A Numerical Algorithm for Zero Counting. III: Randomization and Condition

Felipe Cucker *
Dept. of Mathematics
City University of Hong Kong
HONG KONG
e-mail: macucker@cityu.edu.hk

Teresa Krick †
Departamento de Matemática
Univ. de Buenos Aires & IMAS, CONICET
ARGENTINA
e-mail: krick@dm.uba.ar

Gregorio Malajovich‡
Depto. de Matemática Aplicada
Univ. Federal do Rio de Janeiro
BRASIL
e-mail: gregorio@ufrj.br

Mario Wschebor
Centro de Matemática
Universidad de la República
URUGUAY
e-mail: wschebor@cmat.edu.uy

Abstract. In a recent paper [7] we analyzed a numerical algorithm for computing the number of real zeros of a polynomial system. The analysis relied on a condition number $\kappa(f)$ for the input system f. In this paper we look at $\kappa(f)$ as a random variable derived from imposing a probability measure on the space of polynomial systems and give bounds for both the tail $\mathbb{P}\{\kappa(f) > a\}$ and the expected value $\mathbb{E}(\log \kappa(f))$.

Keywords: Zero-counting, finite-precision, condition numbers, average-case analysis, Rice formula.

Mathematics Subject Classification: 12Y05,60G60,65Y20.

*Partially supported by GRF grant City University 100810.
†Partially supported by grants ANPCyT 33671/05, UBACyT X113/2008-2010 and CONICET PIP/2010-2012.
‡Partially supported by CNPq grants 470031/2007-7, 303565/2007-1, and by FAPERJ.

1 Introduction

1.1 Overview

This paper is the third of a series which started with [7, 8]. In the first paper of the series we analyzed a numerical algorithm for computing the number of real zeros of a polynomial system. This algorithm works with finite precision and the analysis provided bounds for both its complexity (total number of arithmetic operations) and the machine precision needed to guarantee that the returned value is correct. Both bounds depended on size parameters for the input system f (number of polynomials, degrees, etc.) as well as on a condition number $\kappa(f)$ for f. A precise statement of the main result in [7] is Theorem 1.1 therein. To the best of our knowledge, this theorem is the only result providing a finite-precision analysis of a zero counting algorithm. Consequently, as of today, to understand zero-counting computations in the presence of finite-precision appears to require an understanding of $\kappa(f)$.

Unlike the aforementioned size parameters, the condition number $\kappa(f)$ cannot be read directly from the system f. Indeed, it is conjectured that the computation of $\kappa(f)$ is at least as difficult as solving the zero counting problem for f, so we need a much depper understanding of $\kappa(f)$. In the second paper of the series [8], we attempted to provide such an understanding from two different angles. Firstly, we showed that a closely related condition number $\widetilde{\kappa}(f)$ satisfies a Condition Number Theorem, i.e., $\widetilde{\kappa}(f)$ is the normalized inverse of the distance from f to the set of ill-posed systems (those having multiple zeros). The relation between the quantities $\kappa(f)$ and $\widetilde{\kappa}(f)$ is close indeed (see [8, Prop. 3.3]):

$$\frac{\widetilde{\kappa}(f)}{\sqrt{n}} \leq \kappa(f) \leq \sqrt{2n}\, \widetilde{\kappa}(f).$$

Secondly, we used this characterization, in conjunction with a result from [6], to provide a smoothed analysis of $\widetilde{\kappa}(f)$ (and hence, of $\kappa(f)$ as well). A smoothed analysis of the complexity and accuracy for the algorithm in [7] immediately follows. Details about smoothed analyses and distance to ill-posedness can be found in the introduction of [8].

As a consequence of the smoothed analysis of $\widetilde{\kappa}(f)$ one immediately obtains an average-case analysis of this condition number. One is left, however, with the feeling that the bounds thus obtained are far from optimal. Indeed, these bounds follow from a result which is general in two aspects. Firstly, it is a smoothed analysis (of which usual average analysis is just a particular case). Secondly, it is derived from a very general result yielding smoothed analysis bounds for condition numbers satisfying a Condition Number Theorem and stated in terms of some geometric invariants (degree and dimension) of the set of ill-posed inputs. The question of whether a finer average analysis can be obtained by using methods more ad-hoc for the problem at hand naturally poses itself.

In this paper we show that such bounds are possible. Loosely speaking, the average analysis in [8] shows a bound for a typical $\widetilde{\kappa}(f)$ - or $\kappa(f)$ - which is of order $\mathcal{D}^2$ where $\mathcal{D}$ is the Bézout number of f. Here we show that $\sqrt{\mathcal{D}}$ is a more accurate upper-bound. This improvement is meaningful, since $\mathcal{D}$ increases exponentially with n. Our main result implies that if the maximum degree $\mathbf{D}$ remains bounded as n grows, $\mathbb{E}(\ln \kappa(f))$ is bounded from above by a quantity equivalent to $\ln(\mathcal{D}^{1/2})$, which according to the Shub-Smale Theorem, see [19], equals the logarithm of the mathematical expectation of the total number of real roots of the polynomial system. More precisely,

$$\limsup_{n\to\infty} \frac{\mathbb{E}(\ln \kappa(f))}{\ln(\mathcal{D}^{1/2})} \leq 1.$$

No non-trivial lower bound has been obtained for the time being as far as we know.

We next proceed to set up the basic notions and notations enabling us to state the above in more precise terms.

1.2 Basic definitions and main result

For $d \in \mathbb{N}$ we denote by $\mathcal{H}_d$ the subspace of $\mathbb{R}[x_0, \dots, x_n]$ of homogeneous polynomials of degree d and, for $\mathbf{d} := (d_1, \dots, d_n)$, we set $\mathcal{H}_{\mathbf{d}} := \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_n}$. We endow $\mathcal{H}_d$ with the Weyl norm which is defined, for $f \in \mathcal{H}_d$, $f(x) = \sum_{|j|=d} a_j x^j$, by

$$\|f\|_W^2 = \sum_{|j|=d} \frac{a_j^2}{\binom{d}{j}}$$

where $x = (x_0, \dots, x_n)$, $j = (j_0, \dots, j_n)$, $|j| := j_0 + \cdots + j_n$, $x^j = x_0^{j_0} \cdots x_n^{j_n}$ and $\binom{d}{j} := \frac{d!}{j_0! \cdots j_n!}$. We then endow $\mathcal{H}_{\mathbf{d}}$ with the norm given by

$$\|f\| := \max_{1 \le i \le n} \|f_i\|_W.$$

For $f = (f_1, \dots, f_n) \in \mathcal{H}_{\mathbf{d}}$, as in [7], we define the following condition number

$$\kappa(f) = \max_{x \in S^n} \min \left\{ \mu_{\mathrm{norm}}(f, x), \frac{\|f\|}{\|f(x)\|_\infty} \right\}$$

with

$$\mu_{\mathrm{norm}}(f, x) = \sqrt{n}\, \|f\| \, \left\| D_x(f)^{-1} M \right\|.$$

Here

- $D_x(f) = Df(x)|_{T_x S^n}$ is the derivative of f along the unit sphere $S^n \subset \mathbb{R}^{n+1}$ at the point x, a linear operator from the tangent space $T_x(S^n)$ to $\mathbb{R}^n$,
- $M := \begin{bmatrix} \sqrt{d_1} & & \\ & \ddots & \\ & & \sqrt{d_n} \end{bmatrix}$ is the scaling $n \times n$ diagonal matrix with diagonal entries the square roots of the degrees $d_i = \deg(f_i)$,
- the norm $\|D_x(f)^{-1} M\|$ is the spectral norm, i.e., the operator norm $\max\{\|D_x(f)^{-1} M\, y\|_2 ; y \in S^n, y \perp x\}$ with respect to $\| \ \|_2$,
- $\|f(x)\|_\infty = \max_{1 \le i \le n} |f_i(x)|$ denotes as usual the infinity norm.

We next impose the probability measure on $\mathcal{H}_{\mathbf{d}}$ defined by Eric Kostlan [15] and Shub-Smale [19]. This measure assumes the coefficients of the polynomials $f_i = \sum_{|j|=d_i} a_j^{(i)} x^j$ are independent, Gaussian, centered random variables, with variances

$$\mathsf{Var}(a_j^{(i)}) = \binom{d_i}{j}.$$

For this distribution, and for $x, y \in \mathbb{R}^{n+1}, 1 \le i, k \le n$, covariances are given by (see Lemma 2.2 below)

$$\mathbb{E}\big(f_i(x) f_k(y)\big) = \delta_{ik} \langle x, y \rangle^{d_i}$$

where δ_{ik} is the Kronecker symbol.

This probability law is invariant under the action of the orthogonal group and permits to perform the computations below, which appear to be much more complicated under other distributions not sharing this invariance property.

To state our main results a number of quantities will be useful. We use the notation

$$\mathbf{D} := \max_{1 \le i \le n} d_i, \quad \mathcal{D} = \prod_{i=1}^{n} d_i, \quad N := \dim \mathcal{H}_{\mathbf{d}} = \sum_{i=1}^{n} \binom{n + d_i}{n}.$$

We note that $\mathcal{D}$ is the Bézout number of the polynomial system. We may assume here that $d_i \ge 2$ for $1 \le i \le n$ since otherwise we could restrict to a system with fewer equations and unknowns. Notice that $N \le n^{\mathbf{D}+2}$.

We are now ready to state our main result.

Theorem 1.1. *Let the random system f satisfy the conditions of the Shub-Smale model and assume $n \ge 3$. Then,*

(i) *For $a > 4\sqrt{2}\,\mathbf{D}^2 n^{7/2} N^{1/2}$ one has*

$$\mathbb{P}\big(\kappa(f) > a\big) \le K_n \frac{\sqrt{2n}(1 + \ln(a/\sqrt{2n}))^{1/2}}{a},$$

where $K_n := 8\mathbf{D}^2 \mathcal{D}^{1/2} N^{1/2} n^{5/2} + 1$.

(ii)

$$\mathbb{E}(\ln \kappa(f)) \le \ln K_n + (\ln K_n)^{1/2} + (\ln K_n)^{-1/2} + \frac{1}{2} \ln(2n).$$

In fact we are going to prove the corresponding result for the alternative quantity $\widetilde{\kappa}(f)$ already considered in [8], since it will enable us to use $\mathbb{L}^2$ methods, which are more adapted to the type of calculations we will perform. We recall that

$$\widetilde{\kappa}(f) = \frac{\|f\|_W}{\big(\min_{x \in S^n}\{\|D_x(f)^{-1}M\|^{-2} + \|f(x)\|_2^2\}\big)^{1/2}}$$

where $\|f\|_W^2 := \sum_{1 \le i \le n} \|f_i\|_W^2$ is the Weyl norm of the system and $\|f(x)\|^2 := \sum_{1 \le i \le n} f_i(x)^2$ denotes the usual Euclidean norm. As we have already mentioned, we have $\frac{\widetilde{\kappa}(f)}{\sqrt{n}} \le \kappa(f) \le \sqrt{2n}\, \widetilde{\kappa}(f)$. Also, as a consequence of [8, Th. 1.1], $\widetilde{\kappa}(f)$ satisfies $\widetilde{\kappa}(f) \ge 1$ for all $f \in \mathcal{H}_{\mathbf{d}}$.

We will therefore obtain Theorem 1.1 as a direct consequence of the following result.

Theorem 1.2. *Let the random system f satisfy the conditions of the Shub-Smale model and assume $n \ge 3$. Then,*

(i) *For* $a > 4\,\mathbf{D}^2 n^3 N^{1/2}$ *one has*

$$\mathbb{P}\big(\widetilde{\kappa}(f) > a\big) \le K_n \frac{(1+\ln a)^{1/2}}{a}$$

where $K_n := 8\mathbf{D}^2\mathcal{D}^{1/2}\,N^{1/2}n^{5/2} + 1$.

(ii)

$$\mathbb{E}(\ln \widetilde{\kappa}(f)) \le \ln K_n + (\ln K_n)^{1/2} + (\ln K_n)^{-1/2}.$$

Theorem 1.1 follows from $\mathbb{P}\big(\kappa(f) > a\big) \le \mathbb{P}\big(\widetilde{\kappa}(f) > a/\sqrt{2n}\big)$, since $\kappa(f) > a \Rightarrow \widetilde{\kappa}(f) > a/\sqrt{2n}$.

The proof of Theorem 1.2 is given in Section 2. It requires a certain number of auxiliary results. With the aim of isolating (and in this way highlighting) the main ideas, we will postpone the proof of these auxiliary results to Section 3, though stating them as needed in the text. This will be indicated by the symbol $\diamondsuit$ at the end of the statement.

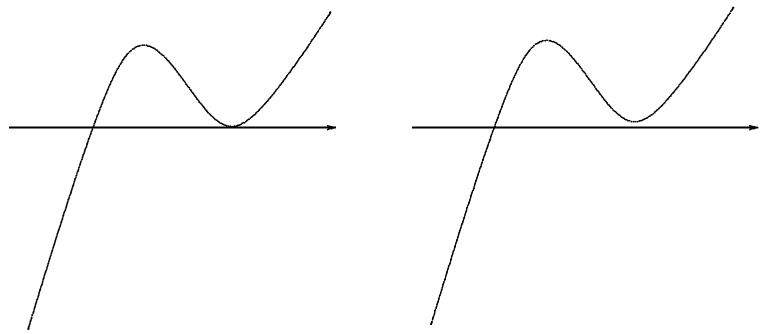
1.3 Relations with previous work

Probably the most successful combination of algorithmics, conditioning, and probability occurs in the study of complex polynomial systems (a setting similar to ours but with the coefficients of the polynomials now drawn from $\mathbb{C}$ and considering projective complex zeros). This study spans an impressive collection of papers, which began with [18, 19, 20, 21, 22] and continued in [3] and [17, 4]. The final outcome of these efforts is a randomized algorithm producing an approximate zero of the input system in expected time which is polynomial in the size of the system. The expectation is with respect to *both* the random choices in the algorithm and a probability measure on the input data.

The condition number of a system f in this setting is defined to be

$$\mu_{\rm norm}(f) := \max_{\zeta\in S^n_{\mathbb{C}}\,|\,f(\zeta)=0} \mu_{\rm norm}(f,\zeta).$$

Here $\mu_{\rm norm}(f,\zeta)$ is roughly the quantity we defined above. Over the reals, it may not be well-defined since the zero set of f may be empty. If one restricts attention to the subset $\mathcal{R}_{\mathbf{d}} \subset \mathcal{H}_{\mathbf{d}}$ of those systems having at least a real zero one may similarly define a measure $\mu_{\rm worst}(f)$, maximizing over the set of real zeros. This has been done in [5] where bounds for the tail and the expected value of $\mu_{\rm worst}(f)$ are given. These bounds are very satisfying (for instance, the tail $\mathbb{P}\big(\mu_{\rm worst} > a\big)$ is bounded by an expression in a^{-2}, a fact ensuring the finiteness of $\mathbb{E}(\mu_{\rm worst}(f))$). The measure $\mu_{\rm worst}(f)$, however, is hardly a condition number for the problem of real zeros counting, not even restricted to the subset $\mathcal{R}_{\mathbf{d}}$. To understand why, consider a polynomial as in the left-hand side of the figure below.



For this polynomial one has $\mu_{\rm worst} = \infty$.

An upward small perturbation (as in the right-hand side) yields a low value of $\mu_{\rm worst}$. This value admits a finite limit when such perturbations are small enough! The measure $\mu_{\rm worst}(f)$ appears to be insensitive to the closeness to ill-posedness. This runs contrary to the notion of conditioning [12, 13, 16, 23].

A condition number $\mu^*(f)$ for the feasibility problem of real systems (which, obviously, needs to be defined on all of $\mathcal{H}_{\mathbf{d}}$) was given in [9] by taking

$$\mu^*(f) = \begin{cases} \min\limits_{\zeta\in S^n | f(\zeta)=0} \mu_{\rm norm}(f,\zeta) & \text{if } f \in \mathcal{R}_{\mathbf{d}} \\ \max\limits_{x\in S^n} \dfrac{\|f\|}{\|f(x)\|} & \text{otherwise.} \end{cases}$$

As of today, there is no probabilistic analysis for it.

2 Proof of Theorem 1.2

The proof relies on the so-called Rice Formula for the expectation of the number of local minima of a real-valued random field. This is described precisely in Step 2 below. Previously, in Step 1, we use large deviations to show that for large n, except on a set of small probability, the numerator $\|f\|_W$ in $\widetilde{\kappa}(f)$ is nearly equal to $N^{1/2}$. Steps 3, 4, and 5 estimate the different expressions occurring in Rice formula. Finally, Step 6 wraps up all these estimates to yield the upper bound for the density and Step 7 derives from it the bounds claimed in the statement of Theorem 1.2.

During the rest of the proof, we set $\underline{L} = \underline{L}(f) := \min_{x\in S^n}\{\|D_x(f)^{-1}M\|^{-2} + \|f(x)\|_2^2\}$ so that $\widetilde{\kappa}(f) = \|f\|_W/\sqrt{\underline{L}}$. We observe that

$$\|D_x(f)^{-1}M\|^{-1} = \sigma_{\min}(M^{-1}D_x(f)) = \min\{\|M^{-1}D_x(f)y\| : y \in S^n, y \perp x\},$$

(where $\sigma_{\min}$ denotes the minimum singular value), and therefore

$$\underline{L} = \min\{\|M^{-1}D_x(f)y\|^2 + \|f(x)\|_2^2 : x, y \in S^n, y \perp x\}$$

is the minimum of the random field $\{L(x,y) : (x,y) \in V\}$ where

$$\begin{aligned} L(x,y) &:= \|M^{-1}D_x(f)y\|^2 + \|f(x)\|_2^2, \\ &= \sum_{i=1}^{n} \frac{1}{d_i}\left(\sum_{j,k=0}^{n} \partial_j f_i(x)\partial_k f_i(x) y_j y_k\right) + \sum_{i=1}^{n} f_i^2(x); \qquad (1) \\ \text{and} \qquad V &:= \{(x,y) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} : \|x\| = \|y\| = 1, \langle x, y\rangle = 0\}. \end{aligned}$$

Here $y = (y_0, \ldots, y_n)$ and, for $1 \le i \le n$ and $0 \le j \le n$, $\partial_j f_i(x)$ denotes the partial derivative of f_i with respect to x_j at the point x.

Step 1. Our first step consists in replacing the Weyl norm in the numerator of $\widetilde{\kappa}(f)$ by a non-random constant, at the cost of adding a small probability, which will be controlled using large deviations.

Let $a > 1$. We have

$$\mathbb{P}\left(\widetilde{\kappa}(f) > a\right) = \mathbb{P}\left(\frac{\underline{L}}{\|f\|_W^2} < \frac{1}{a^2}\right) \le \mathbb{P}\left(\underline{L} < \frac{1}{a^2}(1+\ln a)N\right) + \mathbb{P}\Big(\|f\|_W^2 \ge (1+\ln a)N\Big).$$

We bound the second term in the right-hand side above using the following result that will be proved in Section 3.

Lemma 2.1. *Set*

$$N := \dim \mathcal{H}_{\mathbf{d}} = \sum_{i=1}^{n} \binom{n+d_i}{n}$$

Then, for $\eta > 0$,

$$\mathbb{P}\left(\|f\|_W^2 \ge (1+\eta)N\right) \le e^{-\frac{N}{2}(\eta - \ln(\eta+1))}. \qquad \diamondsuit$$

Therefore, setting $\eta = \ln a$, we obtain

$$\mathbb{P}\left(\widetilde{\kappa}(f) > a\right) \le \mathbb{P}\left(\underline{L} < \frac{1}{a^2}(1+\ln a)N\right) + \exp\left(-\frac{N}{2}(\ln a - \ln(\ln a + 1)\right). \qquad (2)$$

The second term in the right-hand side above can be easily estimated. We therefore turn our attention to the first. Given $\alpha > 0$, we want to compute an upper bound for

$$\mathbb{P}\left(\underline{L} < \alpha\right).$$

Step 2. Our second step consists in giving a bound for the density function $p_{\underline{L}}(u)$ of the random variable $\underline{L}$, i.e. such that

$$\mathbb{P}\left(\underline{L} < \alpha\right) = \int_0^{\alpha} p_{\underline{L}}(u)du$$

since $\underline{L}$ is non-negative. We recall that the quantity $\underline{L}$ is the minimum of the random field $\{L(x,y) : (x,y) \in V\}$, for L and V defined in Formula (1).

Notice that V is the Stiefel manifold $S(2, n+1)$, a compact, orientable, $\mathscr{C}^\infty$-differentiable manifold of dimension $2n-1$, embedded in $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$. For each linear orthogonal transformation U of $\mathbb{R}^{n+1}$, define $\widetilde{U} : V \to V$, $(x, y) \mapsto (Ux, Uy)$, and denote by $\widetilde{\mathcal{U}}$ the set of these $\widetilde{U}$ provided with the group structure naturally inherited from the orthogonal group in $\mathbb{R}^{n+1}$. Then $\widetilde{\mathcal{U}}$ acts transitively on V.

At a generic point (x, y) of the manifold V, the normal space $N_{(x,y)}(V)$ has dimension $(2n+2) - (2n-1) = 3$, and is generated by the orthonormal set $\left\{(x, 0), (0, y), \frac{1}{\sqrt{2}}(y, x)\right\}$. Therefore, if $\{z_2, \dots, z_n\} \subset \mathbb{R}^{n+1}$ is such that $\{x, y, z_2, \dots, z_n\}$ is an orthonormal basis of $\mathbb{R}^{n+1}$, the set

$$\mathcal{B}_{T_{(x,y)}} := \left\{(z_2, 0), \dots, (z_n, 0), (0, z_2), \dots, (0, z_n), \frac{1}{\sqrt{2}}(y, -x)\right\} \tag{3}$$

is an orthonormal basis of the tangent space $T_{(x,y)}(V)$.

We denote by $\sigma_V\big(d(x, y)\big)$ the geometric measure on V (i.e. the measure induced by the Riemannian distance on V), which is invariant under the action of the group $\widetilde{\mathcal{U}}$. The total measure satisfies

$$\sigma_V(V) = \sqrt{2}\sigma_{n-1}\sigma_n, \tag{4}$$

where $\sigma_k = 2\pi^{(k+1)/2}/\Gamma((k+1)/2)$ is the total k-th dimensional measure of the unit sphere S^k, see for example [2, Lemma 13.5].

For $\alpha > 0$ and S a Borel subset of V, we denote by $m_\alpha(L, S)$ the number of local minima of the random function L on the set S, having value smaller than α. Clearly:

$$\mathbb{P}(\underline{L} < \alpha) = \mathbb{P}\big(m_\alpha(L, V) \geq 1\big) \leq \mathbb{E}\big(m_\alpha(L, V)\big). \tag{5}$$

Our aim is to give a useful expression for the right-hand side of Formula (5). For that purpose, let us set for each Borel subset S of V, $\nu(S) := \mathbb{E}\big(m_\alpha(L, S)\big)$. Clearly, ν is a measure. The invariance of the law of the random field $\{L(x, y) : (x, y) \in V\}$ under the action of $\widetilde{\mathcal{U}}$ implies that ν is also invariant under $\widetilde{\mathcal{U}}$.

Let $\psi : B_{2n-1,\delta} \to \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ be a chart on V, that is, a smooth diffeomorphism between the ball in $\mathbb{R}^{2n-1}$ centered at the origin with radius $\delta > 0$ and its image $W = \psi(B_{2n-1,\delta}) \subset V$.

We denote by $\widetilde{L} : B_{2n-1,\delta} \to \mathbb{R}$ the composition $\widetilde{L}(w) = L\big(\psi(w)\big)$.

As we already mentioned, our main tool is Rice formula, of which we now present a quick overview:

Let U be an open subset of $\mathbb{R}^n$ and $Z : U \to \mathbb{R}^n$ a random function having sufficiently smooth paths. Let us denote by $\nu^Z(S)$ the number of zeros of Z belonging to the Borel subset S of U. Under certain general conditions on the probability law of Z, one can compute the expectation of $\nu^Z(S)$ by means of an integral on the set S. The integrand is a certain function depending on the underlying probability law.

The simplest form of such a formula is the following:

$$\mathbb{E}(\nu^Z(S)) = \int_S \mathbb{E}\big(|\det(Z'(t))|/Z(t)=0\big)\, p_{Z(t)}(0)\, dt \tag{6}$$

One must be careful in the choice of the version of the conditional expectation and the density $p_{Z(t)}(\cdot)$ of the random vector $Z(t)$, since they are only defined almost everywhere. But this can be done in a certain number of cases in a canonical form, in such a way that the formula holds true.

This kind of formula can be extended to a variety of situations, such as: a) the zeros of Z can be "marked", which means that instead of all zeros, we count only those zeros satisfying certain additional conditions; b) the domain can be a manifold instead of an open subset of Euclidean space; c) one has formulas similar to (6) for the higher moments of $\nu^Z(S)$; d) the dimension of the domain can be larger than the one of the image, in which case the natural problem, instead of counting roots, is studying the geometry of the random set $Z^{-1}(\{0\})$. For a detailed account of this subject, including proofs and applications, see [2, Chapters 3 and 6].

Here we want to express by means of a Rice formula the expectation

$$\nu(S) = \mathbb{E}\big(m_\alpha(L,S)\big) = \mathbb{E}\big(m_\alpha(\widetilde{L}, \psi^{-1}(S))\big)$$

In our case, with probability 1, $m_\alpha(\widetilde{L}, \psi^{-1}(S))$ equals the number of points $w \in \psi^{-1}(S)$ such that the derivative $\widetilde{L}'(w)$ vanishes, the second derivative $\widetilde{L}''(w)$ is positive definite and the value $L(w)$ is bounded by α. Then, under certain conditions, we can write (use [2, Formula (6.19)], mutatis mutandis):

$$\begin{aligned}\nu(S) &= \mathbb{E}\big(m_\alpha(L,S)\big) = \mathbb{E}\big(m_\alpha(\widetilde{L}, \psi^{-1}(S))\big) \\ &= \int_0^\alpha du \int_{\psi^{-1}(S)} \mathbb{E}\Big(\big|\det(\widetilde{L}''(w))\big| \chi_{\{\widetilde{L}''(w)\succ 0\}} / \widetilde{L}(w) = u, \widetilde{L}'(w) = 0\Big)\, p_{\widetilde{L}(w),\widetilde{L}'(w)}(u,0)\; dw.\end{aligned} \tag{7}$$

Here χ_A means indicator function of the set A, $\succ$ means positive definite, $p_{\widetilde{L}(w),\widetilde{L}'(w)}$ is the joint density in $\mathbb{R}^1 \times \mathbb{R}^{2n-1}$ of the pair of random variables $\big(\widetilde{L}(w), \widetilde{L}'(w)\big)$, and dw is Lebesgue measure on $\mathbb{R}^{2n-1}$. Note that in the chart image, $d\sigma_V = \big(\det\big((\psi'(w))^t\psi'(w)\big)\big)^{1/2} dw$.

In [2, Proposition 6.6] it is proved that if the integrand in Formula (7) were well-defined then the change of variable formula would be satisfied, so that $\nu(S)$ would be the integral of a $(2n-1)$-form. In that case, Formula (7) would already imply that the measure ν is finite and absolutely continuous with respect to σ_V, so that one could write for each Borel subset S of V

$$\nu(S) = \int_S g\; d\sigma_V$$

for a continuous function g. Let us prove that in that case the Radon-Nikodym derivative g would be constant. To see this, notice that σ_V is also invariant under $\widetilde{\mathcal{U}}$ and the action of this group is transitive on V. If g takes different values at two points (x_1, y_1) and (x_2, y_2)

of V, letting $\widetilde{U} \in \widetilde{\mathcal{U}}$ be such that $\widetilde{U}(x_1, y_1) = (x_2, y_2)$, we can find a small neighborhood S of (x_1, y_1) such that

$$\int_S g \, d\sigma_V \neq \int_{\widetilde{U}(S)} g \, d\sigma_V,$$

contradicting the invariance of ν.

We could then compute the constant g by computing it at the point (e_0, e_1). We choose the chart ψ in such a way that $\psi(0) = (e_0, e_1)$ and $\big(\psi'(0)\big)^t \psi'(0) = I_{2n-1}$ and compute

$$\begin{aligned} g &= \lim_{\varepsilon \to 0} \frac{\nu(\psi(B_{2n-1,\varepsilon}))}{\sigma_V(\psi(B_{2n-1,\varepsilon}))} \\ &= \int_0^\alpha \mathbb{E}\left(\big| \det(\widetilde{L}''(0)) \big| \chi_{\{\widetilde{L}''(0) \succ 0\}} / |\widetilde{L}(0) = u, \widetilde{L}'(0) = 0 \right) p_{\widetilde{L}(0), \widetilde{L}'(0)}(u, 0) \, du \end{aligned}$$

So, if Formula (7) were true, it follows that we could write

$$\nu(S) = \sigma_V(S) \int_0^\alpha \mathbb{E}\left(\big| \det(\widetilde{L}''(0)) \big| \chi_{\{\widetilde{L}''(0) \succ 0\}} / \widetilde{L}(0) = u, \widetilde{L}'(0) = 0 \right) p_{\widetilde{L}(0), \widetilde{L}'(0)}(u, 0) \, du. \quad (8)$$

However, if one computes the ingredients in the integrand of the right-hand side of Formula (7), it turns out that the value of the density is $+\infty$ and the conditional expectation vanishes. So, the formula is meaningless in this form.

To overcome this difficulty we proceed as follows:

Let $S_{(x,y)} = \operatorname{span}(z_2, \ldots, z_n) \subset \mathbb{R}^{n+1}$ be the orthogonal complement of $\operatorname{span}(x, y) \subset \mathbb{R}^{n+1}$ and $\pi_{x,y} : \mathbb{R}^{n+1} \to S_{(x,y)}$ be the orthogonal projection. For $(x, y) \in V$, we introduce a new random vector $\zeta_{(x,y)}$ defined as

$$\zeta_{(x,y)} := \Big(\big(\pi_{x,y}(f_i'(x)), \partial_{yy} f_i(x) \big), 1 \le i \le n \Big) \in \big(S_{(x,y)} \times \mathbb{R} \big)^n \cong \mathbb{R}^{n^2}, \quad (9)$$

where for $1 \le i \le n$, $f_i'(x)$ is the free derivative (the gradient) of f_i at x, the first $(n-1)$ coordinates are given by the coordinates of the projection of $f_i'(x)$ onto $S_{(x,y)}$ in the orthonormal basis $\{z_2, \ldots, z_n\}$ and the n-th one is the second derivative in the direction y at x.

Then, instead of Formula (7) we write the formula

$$\begin{aligned} \mathbb{E}\big(m_\alpha(L, S)\big) = \int_0^\alpha du \int_{\psi^{-1}(S)} dw \int_{(S_{\psi(w)} \times \mathbb{R})^n} \mathbb{E}\Big(\big| \det(\widetilde{L}''(w)) \big| \cdot \chi_{\{\widetilde{L}''(w) \succ 0\}} \,/\, \widetilde{L}(w) = u, \\ \widetilde{L}'(w) = 0, \zeta_{\psi(w)} = z \Big) \cdot p_{\widetilde{L}(w), \widetilde{L}'(w), \zeta_{\psi(w)}}(u, 0, z) \, dz. \end{aligned} \quad (10)$$

Formally, Formula (7) is obtained from Formula (10) by integrating in z.

To prove the validity of Formula (10) one could follow exactly the proof of [2, Formula 6.18] if the random field $\{L(x, y) : (x, y) \in V\}$ were Gaussian. This is not our

case. However, it is in fact a simple function of a Gaussian field, namely it is a quadratic form in the coordinates of f and its first derivatives as shown in Formula (1). It is then easy to show that Formula (10) remains true as it is done for the general Rice formulas in [2, Ch. 6, Section 1.4]. This requires proving: (a) the existence and regularity of the density $p_{\widetilde{L}(w),\widetilde{L}'(w),\zeta_{(\psi(w))}}(u,0,z)$ and (b) with probability 1, 0 is a regular value of $\widetilde{L}'(w)$.

(a) is contained below in the present proof (see Step 4). As for (b), once the regularity of this density will be established, it follows in the same way as [2, Proposition 6.5 (a)].

So, using exactly the same arguments leading to Formula (8) we get:

$$\mathbb{E}\big(m_\alpha(L,V)\big) = \sigma_V(V)\int_0^\alpha du \int_{(S_{\psi(0)}\times\mathbb{R})^n} \mathbb{E}\Big(\big|\det(\widetilde{L}''(0))\big|\cdot\chi_{\{\widetilde{L}''(0)\succ 0\}}\,/\,\widetilde{L}(0)=u,\\ \widetilde{L}'(0)=0, \zeta_{\psi(0)}=z\Big)\cdot p_{\widetilde{L}(0),\widetilde{L}'(0),\zeta_{\psi(0)}}(u,0,z)\,dz.$$

Finally, taking into account Inequality (5) we can conclude that:

$$p_{\underline{L}}(u) \le \sigma_V(V)\int_{(S_{\psi(0)}\times\mathbb{R})^n} \mathbb{E}\Big(\big|\det(\widetilde{L}''(0))\big|\cdot\chi_{\{\widetilde{L}''(0)\succ 0\}}\,/\,\widetilde{L}(0)=u,\\ \widetilde{L}'(0)=0, \zeta_{\psi(0)}=z\Big)\cdot p_{\widetilde{L}(0),\widetilde{L}'(0),\zeta_{\psi(0)}}(u,0,z)\,dz. \tag{11}$$

Step 3. For the rest of the proof we fix the following orthonormal basis $\mathcal{B}_T$ (given in (3)) of the tangent space $T := T_{e_0,e_1}$:

$$\mathcal{B}_T = \Big((e_2,0),\ldots,(e_n,0),(0,e_2),\ldots,(0,e_n),\frac{1}{\sqrt{2}}(e_1,-e_0)\Big). \tag{12}$$

Let us recall that in the right-hand side of Inequality (11) the values of $\widetilde{L}(0),\widetilde{L}'(0),\widetilde{L}''(0)$ are computed using a chart ψ of a neighborhood of (e_0,e_1) such that $\psi(0)=(e_0,e_1)$ and the image by ψ' of the canonical basis of $\mathbb{R}^{2n-1}$ is an orthonormal basis of the tangent space T, that we set to be $\mathcal{B}_T$.

We introduce, for $(x,y)\in V$, the gradient $\nabla\widetilde{L}(x,y)$ which is the orthogonal projection of the free derivative $L'(x,y)$ onto the tangent space $T_{(x,y)}$ and is obviously independent of the parametrizations of the manifold V. One can check by means of a direct computation that

$$\nabla\widetilde{L}(e_0,e_1) = \widetilde{L}'(0)\big(\psi'(0)\big)^t.$$

Then, using the change of variables formula for densities and the fact that $\big(\psi'(0)\big)^t\psi'(0) = I_{2n-1}$, we have:

$$p_{\widetilde{L}(0),\widetilde{L}'(0),\zeta_{\psi(0)}}(u,0,z) = p_{L(e_0,e_1),\nabla\widetilde{L}(e_0,e_1),\zeta_{(e_0,e_1)}}(u,0,z).$$

Notation. To simplify notation, from now on we write f_i (resp. $\partial_k f_i$ and $\partial_{k\ell} f_i$, $0\le k,\ell\le n$) for $f_i(e_0)$ (resp. $\partial_k f_i(e_0) = \frac{\partial f_i}{\partial x_k}(e_0)$, $\partial_{k\ell} f_i(e_0) = \frac{\partial^2 f_i}{\partial x_k\partial x_\ell}(e_0)$, $0\le k,\ell\le n$). In

the same spirit we write L for $L(e_0,e_1)=\widetilde{L}(0)$, $\nabla\widetilde{L}$ for $\nabla\widetilde{L}(e_0,e_1)$ and L'' for $L''(e_0,e_1)$. Finally we write ζ for $\zeta(e_0,e_1)$ and S for $S_{(e_0,e_1)}$.

Under this notation, Inequality (11) becomes:

$$p_{\underline{L}}(u)\le \sigma_V(V)\int_{(S\times\mathbb{R})^n}\mathbb{E}\left(\left|\det(\widetilde{L}'')\right|\cdot\chi_{\{\widetilde{L}''\succ 0\}}\Big/ L=u,\nabla\widetilde{L}=0,\zeta=z\right)p_{L,\nabla\widetilde{L},\zeta}(u,0,z)\;dz. \tag{13}$$

According to the definition of $L(x,y)$ in (1) we have

$$L=\sum_{i=1}^n\frac{1}{d_i}(\partial_1 f_i)^2+\sum_{i=1}^n f_i^2, \tag{14}$$

and, from Definition (9),

$$\zeta:=\zeta_{e_0,e_1}=\big((\partial_2 f_i,\dots,\partial_n f_i,\partial_{11}f_i),1\le i\le n\big)\in\mathbb{R}^{n^2}. \tag{15}$$

We also set $[\nabla\widetilde{L}]_{\mathcal{B}_T}:=(\xi_2,\dots,\xi_n,\eta_2,\dots,\eta_n,\varrho)$ for the coordinates of the gradient $\nabla\widetilde{L}$ in the basis $\mathcal{B}_T$.

Using that the (free) partial derivatives of L at (e_0,e_1) are given by

$$\begin{aligned}\frac{\partial L}{\partial x_k}(e_0,e_1)&=\sum_{i=1}^n\frac{2}{d_i}(\partial_{k1}f_i)(\partial_1 f_i)+\sum_{i=1}^n 2f_i(\partial_k f_i)\quad\text{for}\;\;0\le k\le n\\ \frac{\partial L}{\partial y_\ell}(e_0,e_1)&=\sum_{i=1}^n\frac{2}{d_i}(\partial_1 f_i)(\partial_\ell f_i)\quad\text{for}\;\;0\le \ell\le n,\end{aligned}$$

we obtain

$$\begin{aligned}\xi_j&=\langle L'(e_0,e_1),(e_j,0)\rangle=2\sum_{i=1}^n\frac{1}{d_i}(\partial_{1j}f_i)(\partial_1 f_i)+2\sum_{i=1}^n f_i(\partial_j f_i),\quad 2\le j\le n,\\ \eta_j&=\langle L'(e_0,e_1),(0,e_j)\rangle=2\sum_{i=1}^n\frac{1}{d_i}(\partial_1 f_i)(\partial_j f_i),\;\;2\le j\le n,\\ \varrho&=\langle L'(e_0,e_1),2^{-1/2}(e_1,-e_0)\rangle\\ &=\sqrt{2}\Big[\sum_{i=1}^n\frac{1}{d_i}(\partial_1 f_i)(\partial_{11}f_i)+\sum_{i=1}^n f_i(\partial_1 f_i)\Big]-\sqrt{2}\sum_{i=1}^n\frac{1}{d_i}(\partial_0 f_i)(\partial_1 f_i)\\ &=\sqrt{2}\sum_{i=1}^n\frac{1}{d_i}(\partial_1 f_i)(\partial_{11}f_i).\end{aligned} \tag{16}$$

Here, $\langle\ ,\ \rangle$ denotes the usual inner product in $\mathbb{R}^{n+1}\times\mathbb{R}^{n+1}$ and the last equality in (16) follows from the equalities $\partial_0 f_i=d_i f_i$ for $1\le i\le n$ which are easily verified.

Step 4. In this step we focus on the term $p_{L,\nabla\widetilde{L},\zeta}(u,0,z)$ of (13). To this aim we factor this density as

$$p_{L,\nabla\widetilde{L},\zeta}(u,0,z)\;\;=\;\;q_{L,\nabla\widetilde{L}/\zeta=z}(u,0)\cdot p_\zeta(z) \tag{17}$$

where $q_{L,\nabla\widetilde{L}/\zeta=z}(u,0)$ denotes conditional density.

To study the two terms in the right-hand side of (17), we need a lemma containing the ingredients to compute the distributions and conditional expectations appearing in our proof.

Lemma 2.2. *Let $f \in \mathbb{R}[X_0,\dots,X_n]$ be a homogeneous random polynomial of degree d. Assume that f follows the Shub-Smale model for the probability law of the coefficients, i.e. the coefficients of the polynomial $f = \sum_{|j|=d} a_j X^j$ are independent, Gaussian, centered random variables with variances*
$$\mathsf{Var}(a_j) = \binom{d}{j}.$$
Then

- *For $x,y \in \mathbb{R}^{n+1}$, the covariances satisfy*
$$\mathbb{E}\left(f(x)f(y)\right) = \langle x,y\rangle^d \quad \forall\, x,y \in \mathbb{R}^{n+1},$$
where $\langle\ ,\ \rangle$ is the usual inner product in $\mathbb{R}^{n+1}$.

Moreover, if $e_0 := (1,0,\dots,0)$ is the first vector of the canonical basis of $\mathbb{R}^{n+1}$ and we write f (resp. $\partial_k f$ and $\partial_{k\ell} f$, $0 \le k,\ell \le n$) for $f(e_0)$ (resp. $\partial_k f(e_0) = \frac{\partial f}{\partial x_k}(e_0)$, $\partial_{k\ell} f(e_0) = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(e_0)$, $0 \le k,\ell \le n$), we get the following covariances:

- $\mathbb{E}\left(f\partial_k f\right) = \delta_{k0} d$ *for* $0 \le k \le n$.
- $\mathbb{E}\left((\partial_k f)(\partial_{k'} f)\right) = \delta_{kk'}[d + \delta_{k0} d(d-1)]$ *for* $0 \le k,k' \le n$.
- $\mathbb{E}\left(f(\partial_{k\ell} f)\right) = \delta_{k\ell}\delta_{k0} d(d-1)$ *for* $0 \le k,\ell \le n$.
- $\mathbb{E}\left((\partial_{k\ell} f)(\partial_{k'} f)\right) = d(d-1)\left[(d-2)\delta_{\ell 0}\delta_{k0}\delta_{k'0} + \delta_{k0}\delta_{k'\ell} + \delta_{\ell 0}\delta_{kk'}\right]$ *for* $0 \le k,k',\ell \le n$.
- $\mathbb{E}\left((\partial_{k\ell} f)(\partial_{k'\ell'} f)\right) = d(d-1)\Big\{(d-2)(d-3)\delta_{k0}\delta_{\ell 0}\delta_{k'0}\delta_{\ell'0} + (d-2)\big[\delta_{k0}\delta_{k'0}\delta_{\ell\ell'} + \delta_{k'0}\delta_{\ell 0}\delta_{k\ell'} + \delta_{k0}\delta_{\ell'0}\delta_{k'\ell} + \delta_{\ell 0}\delta_{\ell'0}\delta_{kk'}\big] + \delta_{kk'}\delta_{\ell\ell'} + \delta_{k\ell'}\delta_{k'\ell}\Big\}$ *for* $0 \le k,k',\ell,\ell' \le n$. $\diamondsuit$

We proceed with the study of the two terms in the right-hand side of (17).

Computation of $p_\zeta(z)$: By Lemma 2.2, the n^2 coordinates of ζ in (15) are independent Gaussian centered random variables satisfying that $\mathsf{Var}(\partial_k f_i) = d_i$ and $\mathsf{Var}(\partial_{11} f_i) = 2d_i(d_i - 1)$ for $1 \le i \le n$ and $2 \le k \le n$.
Although we are not going to use the exact expression in the sequel, we can immediately deduce for $z = ((z_{i2},\dots,z_{in},z_{i11}), 1 \le i \le n)$ that

$$p_\zeta(z) \;=\; \frac{1}{(2\pi)^{n^2/2}} \frac{1}{\prod_{i=1}^n d_i^{(n-1)/2} \prod_{i=1}^n (2d_i(d_i-1))^{1/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^n \left(\sum_{j=2}^n \frac{z_{ij}^2}{d_i} + \frac{z_{i11}^2}{2d_i(d_i-1)}\right)\right).$$

Computation of $q_{L,\nabla\widetilde{L}/\zeta=z}(0)$: We factor it as follows:
$$q_{L,\nabla\widetilde{L}/\zeta=z}(u,0) = q_{L/\nabla\widetilde{L}=0,\zeta=z}(u)\cdot q_{\nabla\widetilde{L}/\zeta=z}(0).$$

Remembering that $\left(\nabla\widetilde{L}\right)_{\mathcal{B}_T} := (\xi_2,\dots,\xi_n,\eta_2,\dots,\eta_n,\varrho)$, we can write $q_{\nabla\widetilde{L}/\zeta=z}(0)$ as

$$q_{\nabla\widetilde{L}/\zeta=z}(0) = q_{(\xi_2,\dots,\xi_n)/(\eta_2,\dots,\eta_n,\varrho)=0,\ \zeta=z}(0)\cdot q_{(\eta_2,\dots,\eta_n,\varrho)/\zeta=z}(0).$$

First we compute $q_{(\eta_2,\dots,\eta_n,\varrho)/\zeta=z}(0)$. The condition $\zeta = z$ says that for $1\le i\le n$ and $2\le j\le n$, $\partial_j f_i = z_{ij}$ and $\partial_{11} f_i = z_{i11}$. Therefore, from Identities (16), we have

$$\begin{pmatrix}\eta_2\\ \vdots\\ \eta_n\\ \varrho\end{pmatrix} = A(z)\begin{pmatrix}\frac{\partial_1 f_1}{\sqrt{d_1}}\\ \vdots\\ \frac{\partial_1 f_n}{\sqrt{d_n}}\end{pmatrix},\quad\text{where}\quad A(z) = \begin{array}{c} \longleftarrow\ n\ \longrightarrow \\ \left(\begin{array}{ccc}\frac{2}{\sqrt{d_1}}z_{12} & \cdots & \frac{2}{\sqrt{d_n}}z_{n2}\\ \vdots & & \vdots\\ \frac{2}{\sqrt{d_1}}z_{1n} & \cdots & \frac{2}{\sqrt{d_n}}z_{nn}\\ \hline \frac{\sqrt{2}}{\sqrt{d_1}}z_{111} & \cdots & \frac{\sqrt{2}}{\sqrt{d_n}}z_{n11}\end{array}\right)\end{array}\begin{array}{c}\uparrow\\ n-1\\ \downarrow\\ 1\end{array} \tag{18}$$

is non-singular for almost every $z\in\mathbb{R}^{n^2}$. Applying again Lemma 2.2, $\partial_1 f_i/\sqrt{d_i}$, $1\le i\le n$, are independent standard normal random variables that are independent from ζ. By the change of variables formula, we get

$$q_{(\eta_2,\dots,\eta_n,\varrho)/\zeta=z}(0) = \frac{1}{(2\pi)^{n/2}}\cdot\frac{1}{|\det A(z)|}.$$

Now we compute $q_{(\xi_2,\dots,\xi_n)/(\eta_2,\dots,\eta_n,\varrho)=0,\zeta=z}(0)$. Since $A(z)$ is non-singular for almost every z, the condition $\eta_2=\ldots=\eta_n=\varrho=0$ implies $\partial_1 f_i = 0$ for $1\le i\le n$. Therefore, from Identities (16) and since $\zeta = z$, we have

$$\begin{pmatrix}\xi_2\\ \vdots\\ \xi_n\end{pmatrix} = 2\,B(z)\begin{pmatrix}f_1\\ \vdots\\ f_n\end{pmatrix},\quad\text{where}\quad B(z) = \begin{array}{c}\longleftarrow\ n\ \longrightarrow\\ \begin{pmatrix}z_{12} & \cdots & z_{n2}\\ \vdots & & \vdots\\ z_{1n} & \cdots & z_{nn}\end{pmatrix}\end{array}\begin{array}{c}\uparrow\\ n-1\\ \downarrow\end{array}.$$

Again, $f_1,\dots,f_n$ are independent standard normal variables independent from $(\eta_2,\dots,\eta_n,\varrho,\zeta)$ and thus

$$q_{(\xi_2,\dots,\xi_n)/(\eta_2,\dots,\eta_n,\varrho)=0,\ \zeta=z}(0) = \frac{1}{(2\pi)^{(n-1)/2}}\cdot\frac{1}{2^{n-1}(\det(B(z)B(z)^t))^{1/2}},$$

where $B(z)^t$ denotes the transpose of the matrix $B(z)$.
We therefore obtain

$$\begin{aligned} q_{\nabla\widetilde{L}/\zeta=z}(0) &= q_{(\xi_2,\dots,\xi_n)/(\eta_2,\dots,\eta_n,\varrho)=0,\ \zeta=z}(0)\cdot q_{(\eta_2,\dots,\eta_n,\varrho)/\zeta=z}(0)\\ &= \frac{1}{(2\pi)^{n-\frac12}2^{n-1}|\det A(z)|(\det(B(z)B(z)^t))^{1/2}}.\end{aligned}$$

Finally we compute $q_{L/\nabla\widetilde{L}=0,\zeta=z}(u)$. The conditions $\nabla\widetilde{L}=0$ and $\zeta = z$ imply by (18) and (16) that $\partial_1 f_i = 0$ for $1\le i\le n$ and $\sum_{i=1}^n f_i z_{ij} = 0$ for $2\le j\le n$ for almost every z . Plugging the former into (14) we get

$$L = \sum_{i=1}^n f_i^2,$$

and the latter says that the vector $(f_1, \dots, f_n)$ is orthogonal to the $(n-1)$-dimensional subspace S spanned by the $n-1$ vectors $(z_{1j}, \dots, z_{nj})$, $2 \le j \le n$. This shows that $f_1^2 + \cdots + f_n^2$, the square of the distance of $(f_1, \dots, f_n)$ to S, has the χ_1^2-distribution, since the property of being a vector of independent standard normal variables is independent of the choice of the orthonormal basis. So, for $u > 0$,

$$q_{L/\nabla\widetilde{L}=0,\ \zeta=z}(u) = \frac{e^{-u/2}}{\sqrt{2\pi u}}.$$

We therefore obtain

$$\begin{aligned} q_{L,\nabla\widetilde{L}/\zeta=z}(u,0) &= q_{L/\nabla\widetilde{L}=0,\zeta=z}(u) \cdot q_{\nabla\widetilde{L}/\zeta=z}(0) \\ &= \frac{e^{-u/2}}{(2\pi)^n 2^{n-1} |\det\big(A(z)\big)| \big(\det\big(B(z)B(z)^t\big)\big)^{1/2} \sqrt{u}}. \end{aligned}$$

Plugging this expression into Identity (17) we obtain

$$p_{L,\nabla\widetilde{L},\zeta}(u,0,z) = \frac{e^{-u/2}}{(2\pi)^n 2^{n-1} |\det\big(A(z)\big)| \big(\det\big(B(z)B(z)^t\big)\big)^{1/2} \sqrt{u}} \cdot p_\zeta(z). \tag{19}$$

Step 5. In this step we focus on the conditional expectation

$$\mathbb{E}\Big(\big|\det(\widetilde{L}'')\big| \cdot \chi_{\{\widetilde{L}'' \succ 0\}} \Big/ L = u, \nabla\widetilde{L} = 0, \zeta = z\Big) \tag{20}$$

in the integrand of (13). We obtain the following expression for $\widetilde{L}''$ under the stated conditions.

Lemma 2.3. *Let M be the symmetric block-matrix $\mathbb{R}^{(2n-1)\times(2n-1)}$ of the linear operator $\widetilde{L}''$, under the conditions $L = u, \nabla\widetilde{L} = 0$ and $\xi = z$. Let f^* be any solution of the system $\sum_{i=1}^n f_i z_{ij} = 0$, $2 \le j \le n$, and $\sum_{i=1}^n f_i^2 = u$. Then*

$$M = \begin{array}{cc} \begin{array}{ccc} n-1 & n-1 & 1 \end{array} & \\ \left(\begin{array}{c|c|c} M_{\sigma\sigma} & M_{\sigma\tau} & M_{\sigma\theta} \\ \hline M_{\tau\sigma} & M_{\tau\tau} & M_{\tau\theta} \\ \hline M_{\theta\sigma} & M_{\theta\tau} & M_{\theta\theta} \end{array}\right) & \begin{array}{c} n-1 \\ n-1 \\ 1 \end{array} \end{array}$$

where

$$
\begin{aligned}
(M_{\sigma\sigma})_{jj} &= 2\sum_{i=1}^{n}\left(\frac{1}{d_i}(\partial_{1j}f_i)^2 + z_{ij}^2 + f_i^*(\partial_{jj}f_i) - d_i f_i^{*2}\right) \quad \textit{for } 2 \le j \le n,\\
(M_{\sigma\sigma})_{jk} &= 2\sum_{i=1}^{n}\left(\frac{1}{d_i}(\partial_{1j}f_i)(\partial_{1k}f_i) + z_{ij}z_{ik} + f_i^*(\partial_{jk}f_i)\right) \quad \textit{for } 2 \le j \ne k \le n,\\
(M_{\sigma\tau})_{jk} &= 2\sum_{i=1}^{n}\frac{1}{d_i}(\partial_{1j}f_i)z_{ik} \quad \textit{for } 2 \le j,k \le n,\\
(M_{\sigma\theta})_{j1} &= \sqrt{2}\sum_{i=1}^{n}\frac{1}{d_i}(\partial_{1j}f_i)z_{i11} \quad \textit{for } 2 \le j \le n,\\
(M_{\tau\tau})_{jk} &= 2\sum_{i=1}^{n}\frac{1}{d_i}z_{ij}z_{ik} \quad \textit{for } 2 \le j,k \le n,\\
(M_{\tau\theta})_{j1} &= \sqrt{2}\sum_{i=1}^{n}\frac{1}{d_i}z_{i11}z_{ij} \quad \textit{for } 2 \le j \le n,\\
M_{\theta\theta} &= \sum_{i=1}^{n}\left(\frac{1}{d_i}z_{i11}^2 - f_i^* z_{i11}\right).
\end{aligned}
$$

Proof. The hypotheses imply that for almost every z, one has $\partial_1 f_i = 0$ for $1 \le i \le n$, $\sum_{i=1}^n f_i z_{ij} = 0$ for $2 \le j \le n$ and $\sum_{i=1}^n f_i^2 = u$. The last two conditions give a system of n equations and n unknowns with exactly two solutions $f^* = (f_1^*, \ldots, f_n^*)$ and $-f^*$ for almost every z and $u > 0$. Moreover the symmetry of the Gaussian distribution implies that the law of the coordinates of the matrix M does not change under the stated conditions when replacing $f_1, \ldots, f_n$ by either one of these solutions. The formulas are then a consequence of Corollary 3.2 of Section 3 (here we use that $\partial_0 f_i = d_i f_i$ and skip the details).

□

For z fixed, the only random variables that appear in the elements of M are the second partial derivatives $\partial_{jk} f_i$, $2 \le j,k \le n$ and $\partial_{1j} f_i$, $2 \le j \le n, 1 \le i \le n$. Therefore, we are in condition to apply the following result which gets rid of conditioning in (20).

Lemma 2.4. *Let $X = (X_{ij})_{1\le i\le p, 1\le j\le q}$ be a real random matrix and $Y = (Y_1, ..., Y_q)^t$, $Z = (Z_1, ..., Z_p)^t$ be real random vectors. Assume that X, Y, Z are independent, the distributions of X, Y and Z have bounded continuous densities, respectively in $\mathbb{R}^{p\times q}$, $\mathbb{R}^q$, $\mathbb{R}^p$ and that $p_Y(.)$ and $p_Z(.)$ do not vanish. Let $g : \mathbb{R}^{p\times q} \to \mathbb{R}$ be continuous, such that $\mathbb{E}(|g(X)|) < +\infty$. Then, for any $u \in \mathbb{R}^p$,*

$$\mathbb{E}\big(g(X) \,/\, XY + Z = u, Y = 0\big) = \mathbb{E}\big(g(X)\big). \qquad \diamondsuit$$

The heuristic meaning of the previous lemma is that *if we know that* $Y = 0$, then $XY + Z$ does not give information on the distribution of X.

For $X = \left(\frac{1}{d_i}\partial_{1j} f_i\right)_{2\le j\le n, 1\le i\le n} \in \mathbb{R}^{(n-1)\times n}$ and $Y = \left(\partial_1 f_1, \ldots, \partial_1 f_n\right)^t$ in the previous lemma we obtain that

$$\mathbb{E}\left(\left|\det(\widetilde{L}'')\right| \cdot \chi_{\{\widetilde{L}''\succ 0\}} \Big/ L = u, \nabla\widetilde{L} = 0, \zeta = z\right) = \mathbb{E}\left(\left|\det(M)\right| \cdot \chi_{\{M\succ 0\}}\right). \tag{21}$$

We now consider $\mathbb{E}\left(\left|\det(M)\right| \cdot \chi_{\{M \succ 0\}}\right)$. We observe that it is now an unconditional expectation. We will bound it in terms of u and z. We begin by writing the matrix M in a form that will be useful for our computations.

Notation To simplify notation, from now on we simply write A and B for the matrices $A(z)$ and $B(z)$ of Step 4.

We first observe that

$$M_{\sigma\sigma} = VV^t + 2BB^t + W - \mu I_{n-1}$$

where

$$V := \begin{array}{c c} \longleftarrow \quad n \quad \longrightarrow & \\ \begin{pmatrix} \frac{\sqrt{2}}{\sqrt{d_1}}\partial_{12}f_1 & \dots & \frac{\sqrt{2}}{\sqrt{d_n}}\partial_{12}f_n \\ \vdots & & \vdots \\ \frac{\sqrt{2}}{\sqrt{d_1}}\partial_{1n}f_1 & \dots & \frac{\sqrt{2}}{\sqrt{d_n}}\partial_{nn}f_n \end{pmatrix} & \begin{matrix} \uparrow \\ n-1 \\ \downarrow \end{matrix} \end{array}, \quad W := \begin{array}{c c} \longleftarrow \quad n-1 \quad \longrightarrow & \\ \begin{pmatrix} 2\sum_{i=1}^n f_i^*\partial_{22}f_i & \dots & 2\sum_{i=1}^n f_i^*\partial_{2n}f_i \\ \vdots & & \vdots \\ 2\sum_{i=1}^n f_i^*\partial_{n2}f_i & \dots & 2\sum_{i=1}^n f_i^*\partial_{nn}f_i \end{pmatrix} & \begin{matrix} \uparrow \\ n-1 \\ \downarrow \end{matrix} \end{array}$$

and

$$\mu := 2\sum_{i=1}^n d_i {f_i^*}^2.$$

Also, introducing for $1 \le i \le n$ and $2 \le j \le n$,

$$\widetilde{z}_{ij} := \frac{2}{\sqrt{d_i}} z_{ij}, \quad \widetilde{z}_j := (\widetilde{z}_{1j}, \dots, \widetilde{z}_{nj}), \quad \widehat{B} = \widehat{B}(z) := \begin{pmatrix} \widetilde{z}_2 \\ \vdots \\ \widetilde{z}_n \end{pmatrix} = \begin{array}{c c} \longleftarrow \quad n \quad \longrightarrow & \\ \begin{pmatrix} \widetilde{z}_{12} & \dots & \widetilde{z}_{n2} \\ \vdots & & \vdots \\ \widetilde{z}_{1n} & \dots & \widetilde{z}_{nn} \end{pmatrix} & \begin{matrix} \uparrow \\ n-1 \\ \downarrow \end{matrix} \end{array}$$

and

$$\widetilde{z}_{i11} := \frac{2}{\sqrt{d_i}} z_{i11}, \;\; \widetilde{f}_i := \sqrt{d_i} f_i^* \quad \text{and} \quad \widetilde{z}_{11} := (\widetilde{z}_{111}, \dots, \widetilde{z}_{n11}), \;\; \widetilde{f} := (\widetilde{f}_1, \dots, \widetilde{f}_n)$$

so that

$$A = \begin{array}{c c} n & \\ \left(\begin{array}{c} \widehat{B} \\ \hline \frac{1}{\sqrt{2}}\widetilde{z}_{11} \end{array}\right) & \begin{matrix} n-1 \\ \\ 1 \end{matrix} \end{array},$$

we get

$$M_{\sigma\tau} = \frac{1}{\sqrt{2}} V\widehat{B}^t, \; M_{\sigma\theta} = \frac{1}{2} V\widetilde{z}_{11}^t, \; M_{\tau\tau} = \frac{1}{2}\widehat{B}\widehat{B}^t, \; M_{\tau\theta} = \frac{\sqrt{2}}{4}\widehat{B}\widetilde{z}_{11}^t \text{ and } M_{\theta\theta} = \frac{1}{4}\widetilde{z}_{11}\widetilde{z}_{11}^t - \frac{1}{2}\widetilde{z}_{11}\widetilde{f}^t.$$

Therefore

$$M = \begin{array}{c c} \begin{array}{c c c} n-1 & n-1 & 1 \end{array} & \\ \left(\begin{array}{c|c|c} VV^t + 2BB^t + W - \mu I_{n-1} & \frac{1}{\sqrt{2}} V\widehat{B}^t & \frac{1}{2} V\widetilde{z}_{11}^t \\ \hline \frac{1}{\sqrt{2}}\widehat{B}V^t & \frac{1}{2}\widehat{B}\widehat{B}^t & \frac{\sqrt{2}}{4}\widehat{B}\widetilde{z}_{11}^t \\ \hline \frac{1}{2}\widetilde{z}_{11}V^t & \frac{\sqrt{2}}{4}\widetilde{z}_{11}\widehat{B}^t & \frac{1}{4}\widetilde{z}_{11}\widetilde{z}_{11}^t - \frac{1}{2}\widetilde{z}_{11}\widetilde{f}^t \end{array}\right) & \begin{matrix} n-1 \\ n-1 \\ 1 \end{matrix} \end{array}.$$

The coefficients of the matrix W appearing in the first block are the centered Gaussian random variables $\{2\sum_{i=1}^n f_i^* \partial_{jk} f_i \,:\, 2 \le j \le k \le n\}$ which are independent. Applying Lemma 2.2, we obtain

$$\begin{aligned} \sigma^2 := \mathsf{Var}(2\sum_{i=1}^n f_i^* \partial_{jk} f_i) &= 4\sum_{i=1}^n d_i(d_i-1) f_i^{*2} \;\le\; 4\mathbf{D}(\mathbf{D}-1)u \quad \text{for} \quad j \neq k, \\ \mathsf{Var}(2\sum_{i=1}^n f_i^* \partial_{jj} f_i) &= 8\sum_{i=1}^n d_i(d_i-1) f_i^{*2} = 2\sigma^2. \end{aligned} \tag{22}$$

As a consequence, dividing each coefficient of W by $\sigma\sqrt{n-1}$, one can write the matrix W in the form:

$$W = \sigma\sqrt{n-1}\, G$$

where G is a real random symmetric matrix with entries a_{ij} which are independent Gaussian centered satisfying that $\mathsf{Var}(a_{ij}) = 1/n$ for $i \neq j$ and $\mathsf{Var}(a_{ij}) = 2/n$ for $i = j$.

We continue now with the bound for $\mathbb{E}\left(\left|\det(M)\right| \cdot \chi_{\{M \succ 0\}}\right)$. The randomness for this expectation lies in the matrices V and W, which are stochastically independent by Lemma 2.2.

Denote by $\overline{\lambda}$ the maximum between 0 and the largest eigenvalue of the matrix G. Using the independence of V and W, and the fact that the determinant of a positive semidefinite matrix is an increasing function of the diagonal values, we get

$$\mathbb{E}\left(\left|\det(M)\right| \cdot \chi_{\{M \succ 0\}}\right) \;\le\; \mathbb{E}\left(\left|\det(M_1)\right| \cdot \chi_{\{M_1 \succ 0\}}\right) \tag{23}$$

where M_1 is given by:

$$M_1 = \begin{array}{cc} \begin{array}{ccc} {\scriptstyle n-1} & {\scriptstyle n-1} & {\scriptstyle 1} \end{array} & \\ \left(\begin{array}{c|c|c} VV^t + 2BB^t + \sigma\sqrt{n}\,\overline{\lambda}\, I_{n-1} & \frac{1}{\sqrt{2}} V \widehat{B}^t & \frac{1}{2} V \widetilde{z}_{11}^t \\ \hline \frac{1}{\sqrt{2}} \widehat{B} V^t & \frac{1}{2} \widehat{B} \widehat{B}^t & \frac{\sqrt{2}}{4} \widehat{B} \widetilde{z}_{11}^t \\ \hline \frac{1}{2} \widetilde{z}_{11} V^t & \frac{\sqrt{2}}{4} \widetilde{z}_{11} \widehat{B}^t & \frac{1}{4} \widetilde{z}_{11} \widetilde{z}_{11}^t - \frac{1}{2} \widetilde{z}_{11} \widetilde{f}^t \end{array}\right) & \begin{array}{c} {\scriptstyle n-1} \\ {\scriptstyle n-1} \\ {\scriptstyle 1} \end{array} \end{array} .$$

We note that

$$\det(M_1) = \det(M_2) - \frac{1}{2} \widetilde{z}_{11} \widetilde{f}^t \det(M_0). \tag{24}$$

where

$$M_0 = \begin{array}{cc} \begin{array}{cc} {\scriptstyle n-1} & {\scriptstyle n-1} \end{array} & \\ \left(\begin{array}{c|c} VV^t + 2BB^t + \sigma\sqrt{n}\,\overline{\lambda}\, I_{n-1} & \frac{1}{\sqrt{2}} V \widehat{B}^t \\ \hline \frac{1}{\sqrt{2}} \widehat{B} V^t & \frac{1}{2} \widehat{B} \widehat{B}^t \end{array}\right) & \begin{array}{c} {\scriptstyle n-1} \\ {\scriptstyle n-1} \end{array} \end{array} .$$

and

$$M_2 = \begin{array}{cc} \begin{array}{ccc} {\scriptstyle n-1} & {\scriptstyle n-1} & {\scriptstyle 1} \end{array} & \\ \left(\begin{array}{c|c|c} VV^t + 2BB^t + \sigma\sqrt{n}\,\overline{\lambda}\, I_{n-1} & \frac{1}{\sqrt{2}} V \widehat{B}^t & \frac{1}{2} V \widetilde{z}_{11}^t \\ \hline \frac{1}{\sqrt{2}} \widehat{B} V^t & \frac{1}{2} \widehat{B} \widehat{B}^t & \frac{\sqrt{2}}{4} \widehat{B} \widetilde{z}_{11}^t \\ \hline \frac{1}{2} \widetilde{z}_{11} V^t & \frac{\sqrt{2}}{4} \widetilde{z}_{11} \widehat{B}^t & \frac{1}{4} \widetilde{z}_{11} \widetilde{z}_{11}^t \end{array}\right) & \begin{array}{c} {\scriptstyle n-1} \\ {\scriptstyle n-1} \\ {\scriptstyle 1} \end{array} \end{array} .$$

Observe that M_0 and M_2 can be written as

$$M_0 = N_0\, N_0^t \quad \text{and} \quad M_2 = N_2\, N_2^t$$

where

$$N_0 := \begin{array}{cc} \begin{array}{ccc} \scriptstyle n & \scriptstyle n & \scriptstyle n-1 \end{array} & \\ \left(\begin{array}{c|c|c} V & \sqrt{2}B & (\sigma\sqrt{n}\,\overline{\lambda})^{1/2}\, I_{n-1} \\ \hline \frac{1}{\sqrt{2}}\widehat{B} & 0 & 0 \end{array}\right) & \begin{array}{c} \scriptstyle n-1 \\ \scriptstyle n-1 \end{array} \end{array}$$

and

$$N_2 := \begin{array}{cc} \begin{array}{ccc} \scriptstyle n & \scriptstyle n & \scriptstyle n-1 \end{array} & \\ \left(\begin{array}{c|c|c} V & \sqrt{2}B & (\sigma\sqrt{n}\,\overline{\lambda})^{1/2}\, I_{n-1} \\ \hline \frac{1}{\sqrt{2}}\widehat{B} & 0 & 0 \\ \hline \frac{1}{2}\widetilde{z}_{11} & 0 & 0 \end{array}\right) & \begin{array}{c} \scriptstyle n-1 \\ \scriptstyle n-1 \\ \scriptstyle 1 \end{array} \end{array}.$$

Therefore they are both positive semidefinite. Moreover $\det(M_2)$ is the square of the $(2n-1)$-volume of the parallelotope generated by the $2n-1$ rows of N_2. This volume equals the distance from the last row to the subspace generated by the rows of N_0 times the volume of the parallolotope defined by these $2n-2$ rows. The distance from the last row to the subspace generated by the rows of N_0 is bounded by the distance to the smaller subspace generated by the $n-1$ rows of the matrix

$$\left(\begin{array}{c|c|c} \frac{1}{\sqrt{2}}\widehat{B} & 0 & 0 \end{array}\right),$$

which is clearly equal to

$$\operatorname{dist}\left(\frac{1}{2}\widetilde{z}_{11}, \widetilde{S}\right)$$

where $\widetilde{S} := \operatorname{span}(\widetilde{z}_2, \ldots, \widetilde{z}_n) \subset \mathbb{R}^n$. Now we recall that $(f_1^*, \ldots, f_n^*)$ satisfies the conditions $\sum_{i=1}^n f_i^* z_{ij} = 0$, $2 \le j \le n$, which implies

$$\langle \widetilde{f}, \widetilde{z}_j \rangle = 2\sum_{i=1}^n f_i^* z_{ij} = 0,\ \ 2 \le j \le n.$$

This means that $\widetilde{f}$ is orthogonal to $\widetilde{S}$ so that

$$\operatorname{dist}\left(\frac{1}{2}\widetilde{z}_{11}, \widetilde{S}\right) = \frac{1}{2}\left|\left\langle \frac{\widetilde{f}}{\|\widetilde{f}\|}, \widetilde{z}_{11} \right\rangle\right|.$$

Therefore

$$\det(M_2) \le \frac{1}{4}\left|\left\langle \frac{\widetilde{f}}{\|\widetilde{f}\|}, \widetilde{z}_{11} \right\rangle\right|^2 \det(N_0)^2 \ = \ \frac{1}{4}\left|\left\langle \frac{\widetilde{f}}{\|\widetilde{f}\|}, \widetilde{z}_{11} \right\rangle\right|^2 \det(M_0). \tag{25}$$

Using this equality to replace $\det(M_2)$ in (24), we have that

$$|\det(M_1)| \le \frac{1}{2}\left(\frac{1}{2}\left|\left\langle \frac{\widetilde{f}}{\|\widetilde{f}\|}, \widetilde{z}_{11} \right\rangle\right|^2 + \left|\langle \widetilde{f}, \widetilde{z}_{11}\rangle\right|\right) \det(M_0),$$

and therefore, since M_0 is positive semidefinite

$$\mathbb{E}\big(|\det(M_1)|\cdot\chi_{\{M_1\succ 0\}}\big)\le\frac{1}{2}\Big(\frac{1}{2}\Big|\Big\langle\frac{\widetilde{f}}{\|\widetilde{f}\|},\widetilde{z}_{11}\Big\rangle\Big|^2+\big|\langle\widetilde{f},\widetilde{z}_{11}\rangle\big|\Big)\,\mathbb{E}\big(\det(M_0)\big).\qquad(26)$$

We now turn to $\mathbb{E}(\det(M_0))$.

Notation For a matrix M and a subset S (respectively R) of its columns (resp. of its rows), we denote by M^S (resp. M_R) the sub-matrix of M consisting of the columns in S (resp. the rows in R). Also, M_R^S denotes the matrix that consists in erasing the columns not in S and the rows not in R.

Lemma 2.5. *Let $C=(c_{ij})_{i,j}\in\mathbb{R}^{m\times m}$. For $q\in\mathbb{Z}$, $1\le q\le m$, and $\lambda\in\mathbb{R}$ define*

$$C_q(\lambda):=C+\Lambda_q\quad\textit{where}\quad\Lambda_q:=\begin{array}{c c} \begin{array}{cc} \scriptstyle q & \scriptstyle m-q \end{array} & \\ \left(\begin{array}{c|c}\lambda\,\mathrm{Id} & 0\\ \hline 0 & \mathrm{Id}\end{array}\right) & \begin{array}{c}\scriptstyle q\\ \scriptstyle m-q\end{array}\end{array},$$

i.e the matrix obtained by adding λ to the first q diagonal entries of C.
Then,

$$\det(C_q(\lambda))=\det(C)+\sum_{\ell=1}^{q}\Big(\sum_{S\subset\{1,\dots,q\}:\#(S)=\ell}\det\big(C_{\overline{S}}^{\overline{S}}\big)\Big)\lambda^\ell.$$

where $\overline{S}$ is the complement set of S, with the convention that $\det(C_\emptyset^\emptyset)=1$. $\diamondsuit$

We set $\lambda:=\sigma\sqrt{n}\,\overline{\lambda}$ and write $M_0=C+\Lambda$ where

$$C:=\begin{array}{c c} \begin{array}{cc} \scriptstyle n-1 & \scriptstyle n-1 \end{array} & \\ \left(\begin{array}{c|c} VV^t+2BB^t & \frac{1}{\sqrt{2}}V\widehat{B}^t\\ \hline \frac{1}{\sqrt{2}}\widehat{B}V^t & \frac{1}{2}\widehat{B}\widehat{B}^t\end{array}\right) & \begin{array}{c}\scriptstyle n-1\\ \scriptstyle n-1\end{array}\end{array}\quad\text{and}\quad\Lambda:=\begin{array}{c c} \begin{array}{cc} \scriptstyle n-1 & \scriptstyle n-1 \end{array} & \\ \left(\begin{array}{c|c}\lambda\,\mathrm{Id} & 0\\ \hline 0 & \mathrm{Id}\end{array}\right) & \begin{array}{c}\scriptstyle n-1\\ \scriptstyle n-1\end{array}\end{array}.$$

Then, by Lemma 2.5 and using that the random variables involved in the expectation of M_0 are the elements of V and $\overline{\lambda}$, which are independent, we obtain

$$\mathbb{E}\big(\det(M_0)\big)\;=\;\mathbb{E}\big(\det(C)\big)+\sum_{\ell=1}^{n-1}\sum_{\substack{S\subset\{1,\dots,n-1\}\\ \#(S)=\ell}}\mathbb{E}\big(\det\big(C_{\overline{S}}^{\overline{S}}\big)\big)(\sigma\sqrt{n})^\ell\,\mathbb{E}\big(\overline{\lambda}^\ell\big).\qquad(27)$$

We now bound the expectations appearing here. We first consider $\mathbb{E}\big(\det(C)\big)$.

Lemma 2.6. *Set $n,k\in\mathbb{N}$, $1\le k<n$. Let $A=(a_{ij})_{i,j},B\in\mathbb{R}^{k\times n}$ and $C\in\mathbb{R}^{(n-1)\times n}$. Define*

$$Q:=\begin{array}{c c} \begin{array}{cc} \scriptstyle k & \scriptstyle n-1 \end{array} & \\ \left(\begin{array}{c|c} A\,A^t+B\,B^t & A\,C^t\\ \hline C\,A^t & C\,C^t\end{array}\right) & \begin{array}{c}\scriptstyle k\\ \scriptstyle n-1\end{array}\end{array}\in\mathbb{R}^{(k+n-1)\times(k+n-1)}.$$

Then,

$$\det(Q) \;=\; \det(CC^t)\det(BB^t) \;+ \sum_{\#(S)=k-1}\Big(\sum_{i=1}^{k}\sum_{j=1}^{n}(-1)^{i+j-1}a_{ij}\det(B_{\bar i}^{S})\det(C^{\bar j})\Big)^2. \qquad \diamondsuit$$

Applying this result for $k := n-1$, $A := V$, $B := \sqrt{2}B$ and $C := (1/\sqrt{2})\widehat{B}$ we get

$$\det(C) = \det(BB^t)\det(\widehat{B}\widehat{B}^t) + \sum_{\#(S)=n-2}\Big(\sum_{i=1}^{n-1}\sum_{j=1}^{n}(-1)^{i+j-1}v_{ij}\det\big(\sqrt{2}B_{\bar i}^{S}\big)\det\big(\frac{1}{\sqrt{2}}\widehat{B}^{\bar j}\big)\Big)^2.$$

Since the random variables $v_{ij} = \sqrt{2/d_j}\,\partial_{1(i+1)}f_j$ are centered and independent, and since $\mathsf{Var}(v_{ij}) = 2(d_j-1)$, we obtain

$$\begin{aligned}
\mathbb{E}\big(\det(C)\big) &= \det(BB^t)\det(\widehat{B}\widehat{B}^t) + \sum_{\#(S)=n-2}\mathbb{E}\left(\Big(\sum_{i=1}^{n-1}\sum_{j=1}^{n}\pm v_{ij}\det\big(\sqrt{2}B_{\bar i}^{S}\big)\det\big(\frac{1}{\sqrt{2}}\widehat{B}^{\bar j}\big)\Big)^2\right)\\
&= \det(BB^t)\det(\widehat{B}\widehat{B}^t) + \sum_{\#(S)=n-2}\sum_{i=1}^{n-1}\sum_{j=1}^{n}2(d_j-1)2^{n-2}\Big(\det\big(B_{\bar i}^{S}\big)\Big)^2\frac{1}{2^{n-1}}\Big(\det\big(\widehat{B}^{\bar j}\big)\Big)^2\\
&\le \det(BB^t)\det(\widehat{B}\widehat{B}^t) + (\mathbf{D}-1)\sum_{\#(S)=n-2}\sum_{i=1}^{n-1}\sum_{j=1}^{n}\Big(\det\big(B_{\bar i}^{S}\big)\Big)^2\Big(\det\big(\widehat{B}^{\bar j}\big)\Big)^2\\
&= \det(\widehat{B}\widehat{B}^t)\Big(\det(BB^t) + (\mathbf{D}-1)\sum_{i=1}^{n-1}\det\big(B_{\bar i}B_{\bar i}^t\big)\Big)
\end{aligned} \tag{28}$$

where in the last equality we applied twice the well-known Cauchy-Binet formula, see for example [14]: For $m \le n$, $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times m}$,

$$\det(A\,B) = \sum_{S:\#(S)=m}\det(A^S)\det(B_S). \tag{29}$$

Now we compute $\mathbb{E}\big(\det\big(C_{\overline{S}}^{\overline{S}}\big)\big)$ for $\#(S)=\ell$, $1\le \ell\le n-1$.

- For $\ell = n-1$ it is obvious that

$$\det\big(C_{\overline{S}}^{\overline{S}}\big) = (1/2^{n-1})\,\det(\widehat{B}\widehat{B}^t). \tag{30}$$

- For $1\le\ell\le n-2$, we note that for each $S\subset\{1,\dots,n-1\}$ with $\#(S)=\ell$, we have

$$C_{\overline{S}}^{\overline{S}} := \begin{array}{cc} \begin{array}{cc} \scriptstyle n-1-\ell & \scriptstyle n-1 \end{array} & \\ \left(\begin{array}{c|c} V_{\overline{S}}(V_{\overline{S}})^t + 2B_{\overline{S}}(B_{\overline{S}})^t & \frac{1}{\sqrt{2}}V_{\overline{S}}\widehat{B}^t \\ \hline \frac{1}{\sqrt{2}}\widehat{B}(V_{\overline{S}})^t & \frac{1}{2}\widehat{B}\widehat{B}^t \end{array}\right) & \begin{array}{c} \scriptstyle n-1-\ell \\ \scriptstyle n-1 \end{array} \end{array}$$

and we obtain, imitating the computation for the case $\det(C)$,

$$\mathbb{E}\big(\det(C_{\overline{S}}^{\overline{S}})\big) \le \frac{\det(\widehat{B}\widehat{B}^t)}{2^{\ell}}\Big(\det(B_{\overline{S}}(B_{\overline{S}})^t) + (\mathbf{D}-1)\sum_{\substack{1\,\le\, i\,\le\, n-1-\ell\\ i\,\notin\, S}}\det\big(B_{\overline{S\cup\{i\}}}(B_{\overline{S\cup\{i\}}})^t\big)\Big). \tag{31}$$

Finally we give an upper-bound for $\mathbb{E}\big(\overline{\lambda}^{\ell}\big)$.

Lemma 2.7. *Let $G = (a_{ij})_{1\le i,j\le n}$ for $n \ge 2$ be a real random symmetric matrix such that the the random variables $\{a_{ij}, 1 \le i \le j \le n\}$ are independent Gaussian centered, $Var(a_{ij}) = 1/n$ for $i \ne j$ and $Var(a_{ij}) = 2/n$ for $i = j$, and denote by $\overline{\lambda}$ the maximum between 0 and the largest eigenvalue of the matrix G. Then, for $1 \le \ell \le n$,*

$$\mathbb{E}\big(\overline{\lambda}^{\ell}\big) \le 2\cdot 4^{\ell}. \qquad \diamondsuit$$

Plugging Inequalities (28), (31), (30) and Lemma (2.7) into Formula (27) we obtain

$$\begin{aligned}
\mathbb{E}\big(\det(M_0)\big) \le & \det(\widehat{B}\widehat{B}^t)\Big(\det(BB^t) + (\mathbf{D}-1)\sum_{i=1}^{n-1}\det\big(B_{\bar{i}}B_{\bar{i}}^t\big) + 2^n(\sigma\sqrt{n})^{n-1} \\
& + \sum_{\ell=1}^{n-2} 2^{\ell+1}(\sigma\sqrt{n})^{\ell}\Big(\sum_{\#(S)=\ell}\Big(\det(B_{\overline{S}}(B_{\overline{S}})^t) + (\mathbf{D}-1)\sum_{i=1}^{n-1-\ell}\det\big(B_{\overline{S\cup\{i\}}}(B_{\overline{S\cup\{i\}}})^t\big)\Big)\Big)\Big) \\
\le & \det(\widehat{B}\widehat{B}^t)\Big(\det(BB^t) + (\mathbf{D}-1)\sum_{i=1}^{n-1}\det\big(B_{\bar{i}}B_{\bar{i}}^t\big) + 2^n(\sigma\sqrt{n})^{n-1} \\
& + \sum_{\ell=1}^{n-2} 2^{\ell+1}(\sigma\sqrt{n})^{\ell}\Big(\sum_{\#(S)=\ell}\det(B_{\overline{S}}(B_{\overline{S}})^t) + (\mathbf{D}-1)(\ell+1)\sum_{\#(T)=\ell+1}\det\big(B_{\overline{T}}(B_{\overline{T}})^t\big)\Big)\Big).
\end{aligned}$$

This finally implies, by Identity (21) and Inequalities (23) and (26) the inequality we will focuse on in next step.

$$\begin{aligned}
\mathbb{E}\Big(\big|\det(\widetilde{L}'')\big|\cdot\chi_{\{\widetilde{L}''\succ 0\}}\Big/ L = u, \nabla\widetilde{L} = 0, \zeta = z\Big) \le & \frac{1}{2}\Big(\frac{1}{2}\Big|\Big\langle\frac{\widetilde{f}}{\|\widetilde{f}\|}, \widetilde{z}_{11}\Big\rangle\Big|^2 + \big|\langle\widetilde{f}, \widetilde{z}_{11}\rangle\big|\Big)\det(\widehat{B}\widehat{B}^t)\cdot \\
\cdot\Big(\det(BB^t) + (\mathbf{D}-1)\sum_{i=1}^{n-1}\det\big(B_{\bar{i}}B_{\bar{i}}^t\big) + 2^n(\sigma\sqrt{n})^{n-1} & \\
+ \sum_{\ell=1}^{n-2} 2^{\ell+1}(\sigma\sqrt{n})^{\ell}\Big(\sum_{\#(S)=\ell}\det(B_{\overline{S}}(B_{\overline{S}})^t) + (\mathbf{D}-1)(\ell+1)\sum_{\#(T)=\ell+1}\det\big(B_{\overline{T}}(B_{\overline{T}})^t\big)\Big)\Big). & \qquad (32)
\end{aligned}$$

Step 6. We put together the calculations of Steps 4 and 5 to compute an upper bound for $p_{\underline{L}}(u)$ following Inequality (13). We will also use the following auxiliary result:

Lemma 2.8.

$$\frac{2^{2n-1}}{\mathcal{D}}\det(BB^t) \le \det(\widehat{B}\widehat{B}^t) \le \frac{2^{2(n-1)}\mathbf{D}}{\mathcal{D}}\det(BB^t),$$
$$\frac{2^{2n-1-\ell}}{\mathcal{D}}\det(B_{\overline{S}}(B_{\overline{S}})^t) \le \det(\widehat{B_{\overline{S}}}(\widehat{B}_{\overline{S}})^t) \le \frac{2^{2(n-1-\ell)}\mathbf{D}^{\ell+1}}{\mathcal{D}}\det(B_{\overline{S}}(B_{\overline{S}})^t) \quad \textit{for } S\subset\{1,\dots,n\}, \#(S)=\ell.$$

Proof. We have $\widehat{B} = B\ H$ for the diagonal matrix

$$H := \begin{array}{c} \longleftarrow \quad n \quad \longrightarrow \\ \begin{pmatrix} \frac{2}{\sqrt{d_1}} & & \\ & \ddots & \\ & & \frac{2}{\sqrt{d_n}} \end{pmatrix} \end{array} \begin{array}{c} \uparrow \\ n \\ \downarrow \end{array}.$$

By Cauchy-Binet formula (29),

$$\begin{aligned} \det(\widehat{B}\widehat{B}^t) &= \sum_{k=1}^{n} \det(\widehat{B}^{\overline{k}}) \det((\widehat{B}^{\overline{k}})^t) \\ &= \sum_{k=1}^{n} \Big(\det(B^{\overline{k}} H^{\overline{k}}_{\overline{k}}) \Big)^2 = \sum_{k=1}^{n} \big(\det(H^{\overline{k}}_{\overline{k}}) \big)^2 \big(\det(B^{\overline{k}}) \big)^2 \\ &= \sum_{k=1}^{n} \frac{2^{2(n-1)} d_k}{\mathcal{D}} \big(\det(B^{\overline{k}}) \big)^2. \end{aligned}$$

The proof concludes using

$$\frac{2^{2n-1}}{\mathcal{D}} \le \frac{2^{2(n-1)} d_k}{\mathcal{D}} \le \frac{2^{2(n-1)} \mathbf{D}}{\mathcal{D}} \quad \text{since} \quad d_k \ge 2 \quad \text{and} \quad \sum_{k=1}^{n} \big(\det(B^{\overline{k}}) \big)^2 = \det(BB^t).$$

The proof of the second assertion is analogous. □

According to Inequalities (13), (32), and Identity (19), we get:

$$\begin{aligned} p_{\underline{L}}(u) \le \sigma_V(V) &\int_{(S\times\mathbb{R})^n} \mathbb{E}\Big(\big| \det(\widetilde{L}'') \big| \cdot \chi_{\{\widetilde{L}'' \succ 0\}} \Big/ L = u, \nabla \widetilde{L} = 0, \zeta = z \Big) \cdot p_{L,\nabla\widetilde{L},\zeta}(u, 0, z)\ dz \\ \le \sigma_V(V) &\int_{(S\times\mathbb{R})^n} \frac{1}{2}\Big(\frac{1}{2} \Big| \Big\langle \frac{\widetilde{f}}{\|\widetilde{f}\|}, \widetilde{z}_{11} \Big\rangle \Big|^2 + \big| \langle \widetilde{f}, \widetilde{z}_{11} \rangle \big| \Big) \det(\widehat{B}\widehat{B}^t) \cdot \\ &\cdot \Big(\det(BB^t) + (\mathbf{D}-1) \sum_{i=1}^{n-1} \det\big(B_{\overline{i}} B^t_{\overline{i}} \big) + 2^n (\sigma\sqrt{n})^{n-1} \\ &+ \sum_{\ell=1}^{n-2} 2^{\ell+1} (\sigma\sqrt{n})^{\ell} \Big(\sum_{\#(S)=\ell} \det(B_{\overline{S}}(B_{\overline{S}})^t) + (\mathbf{D}-1)(\ell+1) \sum_{\#(T)=\ell+1} \det\big(B_{\overline{T}} (B_{\overline{T}})^t \big) \Big) \Big) \cdot \\ &\frac{e^{-u/2}}{(2\pi)^n 2^{n-1} |\det(A)| \big(\det(BB^t) \big)^{1/2} \sqrt{u}} \cdot p_\zeta(z) dz. \end{aligned}$$

Here we notice that $|\det(A)|$ is the n-volume of the parallelotope generated in $\mathbb{R}^n$ by the rows of A, that is, in the same way we computed $\det(M_2)$ in (25), we have

$$|\det(A)| = \operatorname{dist}\left(\frac{1}{\sqrt{2}} \widetilde{z}_{11}, \widetilde{S} \right) \det(\widehat{B}\widehat{B}^t)^{1/2} = \frac{1}{\sqrt{2}} \Big| \Big\langle \frac{\widetilde{f}}{\|\widetilde{f}\|}, \widetilde{z}_{11} \Big\rangle \Big| \big(\det(\widehat{B}\widehat{B}^t) \big)^{1/2}$$

where like previously $\widetilde{S} := \operatorname{span}(\widetilde{z}_2, \ldots, \widetilde{z}_n) \subset \mathbb{R}^n$ is the hyperplane spanned by the the rows of $\widehat{B}$. Therefore, using Cauchy-Schwartz inequality for $\langle \widetilde{f}/\|\widetilde{f}\|, \widetilde{z}_{11}\rangle$, applying Lemma 2.8 and the fact that $2^n \le \mathcal{D}$, we get

$$
\begin{aligned}
p_{\underline{L}}(u) \le \sigma_V(V) \int_{(S\times\mathbb{R})^n} & \frac{\sqrt{2}}{(4\pi)^n}\big(\frac{1}{2}\|\widetilde{z}_{11}\| + \|\widetilde{f}\|\big)\Big(\frac{\det(\widehat{B}\widehat{B}^t)}{\det(BB^t)}\Big)^{1/2} \cdot \\
& \cdot \Big(\det(BB^t) + (\mathbf{D}-1)\sum_{i=1}^{n-1} \det\big(B_{\overline{i}}B_{\overline{i}}^t\big) + 2^n(\sigma\sqrt{n})^{n-1} \\
& + \sum_{\ell=1}^{n-2} 2^{\ell+1}(\sigma\sqrt{n})^\ell \Big(\sum_{\#(S)=\ell} \det(B_{\overline{S}}(B_{\overline{S}})^t) + (\mathbf{D}-1)(\ell+1)\sum_{\#(T)=\ell+1} \det\big(B_{\overline{T}}(B_{\overline{T}})^t\big)\Big)\Big) \cdot \frac{e^{-u/2}}{\sqrt{u}} p_\zeta(z)dz \\
& \le \sigma_V(V) \int_{(S\times\mathbb{R})^n} \frac{\sqrt{2}}{(4\pi)^n}\big(\frac{1}{2}\|\widetilde{z}_{11}\| + \|\widetilde{f}\|\big) 2^{n-1}\frac{\sqrt{\mathbf{D}}}{\sqrt{\mathcal{D}}} \cdot \\
& \cdot \Big(\frac{\mathcal{D}}{2^{2n-1}}\det(\widehat{B}\widehat{B}^t) + (\mathbf{D}-1)\sum_{i=1}^{n-1}\frac{\mathcal{D}}{2^{2n-2}}\det(\widehat{B_{\overline{i}}}(\widehat{B}_{\overline{i}})^t) + \mathcal{D}(\sigma\sqrt{n})^{n-1} \\
& + \sum_{\ell=1}^{n-2} 2^{\ell+1}(\sigma\sqrt{n})^\ell \Big(\sum_{\#(S)=\ell} \frac{\mathcal{D}}{2^{2n-1-\ell}}\det(\widehat{B}_{\overline{S}}(\widehat{B}_{\overline{S}})^t) + (\mathbf{D}-1)(\ell+1)\sum_{\#(T)=\ell+1}\frac{\mathcal{D}}{2^{2n-\ell}}\det\big(\widehat{B}_{\overline{T}}(\widehat{B}_{\overline{T}})^t\big)\Big)\Big) \cdot \\
& \cdot \frac{e^{-u/2}}{\sqrt{u}} p_\zeta(z)dz \\
& \le \sigma_V(V) \int_{(S\times\mathbb{R})^n} \frac{\sqrt{2}}{(8\pi)^n}\big(\frac{1}{2}\|\widetilde{z}_{11}\| + \|\widetilde{f}\|\big)\sqrt{\mathbf{D}\mathcal{D}} \cdot \Big(\det(\widehat{B}\widehat{B}^t) + 2(\mathbf{D}-1)\sum_{i=1}^{n-1}\det(\widehat{B_{\overline{i}}}(\widehat{B}_{\overline{i}})^t) + 2(4\sigma\sqrt{n})^{n-1} \\
& + \sum_{\ell=1}^{n-2}(4\sigma\sqrt{n})^\ell\Big(\sum_{\#(S)=\ell} 2\det(\widehat{B}_{\overline{S}}(\widehat{B}_{\overline{S}})^t) + (\mathbf{D}-1)(\ell+1)\sum_{\#(T)=\ell+1}\det\big(\widehat{B}_{\overline{T}}(\widehat{B}_{\overline{T}})^t\big)\Big)\Big) \cdot \frac{e^{-u/2}}{\sqrt{u}} p_\zeta(z)dz \\
& = \mathbb{E}\big(H(u,\zeta)\big),
\end{aligned}
$$

where

$$
\begin{aligned}
H(u,\zeta) := & \frac{\sqrt{2}}{(8\pi)^n}\sigma_V(V)\big(\frac{1}{2}\|\widetilde{\zeta}_{11}\| + \|\widetilde{f}\|\big)\sqrt{\mathbf{D}\mathcal{D}} \cdot \Big(\det(\widehat{B}(\zeta)\widehat{B}^t(\zeta)) + 2(4\sigma\sqrt{n})^{n-1} \\
& + 2(\mathbf{D}-1)\sum_{i=1}^{n-1}\det(\widehat{B}_{\overline{i}}(\zeta)(\widehat{B}_{\overline{i}})^t(\zeta)) + \sum_{\ell=1}^{n-2}(4\sigma\sqrt{n})^\ell\Big(\sum_{\#(S)=\ell} 2\det(\widehat{B}_{\overline{S}}(\zeta)(\widehat{B}_{\overline{S}})^t(\zeta)) \\
& + (\mathbf{D}-1)(\ell+1)\sum_{\#(T)=\ell+1}\det\big(\widehat{B}_{\overline{T}}(\zeta)(\widehat{B}_{\overline{T}}(\zeta))^t\big)\Big)\Big) \cdot \frac{e^{-u/2}}{\sqrt{u}}.
\end{aligned}
$$

Here

$$\widehat{B}(\zeta) := \begin{array}{c} \longleftarrow \quad n \quad \longrightarrow \\ \begin{pmatrix} \frac{2}{\sqrt{d_1}}\partial_2 f_1 & \dots & \frac{2}{\sqrt{d_n}}\partial_2 f_n \\ \vdots & & \vdots \\ \frac{2}{\sqrt{d_1}}\partial_n f_1 & \dots & \frac{2}{\sqrt{d_n}}\partial_n f_n \end{pmatrix} \end{array} \begin{array}{c} \uparrow \\ {\scriptstyle n-1} \\ \downarrow \end{array} \quad \text{and} \quad \widetilde{\zeta}_{11} := \big(\frac{2}{\sqrt{d_1}}\partial_{11} f_1, \dots, \frac{2}{\sqrt{d_n}}\partial_{11} f_n\big).$$

Our next goal is then to bound $\mathbb{E}(H(u,\zeta))$. We first note that the matrix $\widehat{B}(\zeta)$ is independent from $\widetilde{\zeta}_{11}$, so that the expectation can be factorized as a product of expectations.

First, using Lemma 2.2 and the definition of $\widetilde{f}$ we easily get

$$\mathbb{E}\big(\frac{1}{2}\|\widetilde{\zeta}_{11}\| + \|\widetilde{f}\|\big) \leq \sqrt{2(\mathbf{D}-1)n} + \sqrt{\mathbf{D}u}.$$

For the other expectations we apply the following.

Lemma 2.9. *(e.g. [2, Lemma 13.6]) Set $m \leq n$ and let U be an $m \times n$ random matrix whose elements are independent real standard normal. Then*

$$\mathbb{E}\big(\det(UU^t)\big) = \frac{n!}{(n-m)!}. \qquad \square$$

Therefore, since by Lemma 2.2, $\frac{1}{2}\widehat{B}(\zeta)$ satisfies the hypothesis of the lemma with $m = n-1$, we obtain

$$\mathbb{E}\big(\det(\widehat{B}(\zeta)\widehat{B}^t(\zeta)\big) = 4^{n-1}n!$$

and we get similar expressions for the other determinants in $\mathbb{E}(H(u,\zeta)))$:

$$\mathbb{E}\big(\det(\widehat{B}_{\overline{i}}(\zeta)(\widehat{B}_{\overline{i}})^t(\zeta))\big) = 4^{n-2}\frac{n!}{2},$$

$$\mathbb{E}\big(\det(\widehat{B}_{\overline{S}}(\zeta)\widehat{B}^t_{\overline{S}}(\zeta))\big) = 4^{n-1-\ell}\frac{n!}{(\ell+1)!},$$

$$\mathbb{E}\big(\det\big(\widehat{B}_{\overline{T}}(\zeta)(\widehat{B}_{\overline{T}}(\zeta))^t\big) = 4^{n-2-\ell}\frac{n!}{(\ell+2)!}.$$

We also apply Formula (4): $\sigma_V(V) = 4\sqrt{2}\pi^{n+\frac{1}{2}}/\big(\Gamma(n/2)\Gamma((n+1)/2)\big)$. Therefore

$$
\begin{aligned}
\mathbb{E}\big(H(u,\zeta)\big) &= \frac{\sqrt{2}}{(8\pi)^n}\frac{4\sqrt{2}\pi^{n+\frac{1}{2}}}{\Gamma(n/2)\Gamma((n+1)/2)}\big(\sqrt{2(\mathbf{D}-1)n}+\sqrt{\mathbf{D}u}\big)\sqrt{\mathbf{D}\mathcal{D}} \\
&\cdot\Bigg(4^{n-1}n!+2(4\sigma\sqrt{n})^{n-1}+2(\mathbf{D}-1)\sum_{i=1}^{n-1}4^{n-2}\frac{n!}{2} \\
&+\sum_{\ell=1}^{n-2}(4\sigma\sqrt{n})^{\ell}\Big(\binom{n-1}{\ell}2\cdot 4^{n-1-\ell}\frac{n!}{(\ell+1)!}+(\mathbf{D}-1)(\ell+1)\binom{n-1}{\ell+1}4^{n-2-\ell}\frac{n!}{(\ell+2)!}\Big)\Bigg) \\
&\cdot\frac{e^{-u/2}}{\sqrt{u}} \\
&= \frac{\sqrt{\pi}}{8^{n-1}\Gamma(n/2)\Gamma((n+1)/2)}\big(\sqrt{2(\mathbf{D}-1)n}+\sqrt{\mathbf{D}u}\big)\sqrt{\mathbf{D}\mathcal{D}}\,4^{n-1}n! \\
&\cdot\Bigg(1+2\frac{(\sigma\sqrt{n})^{n-1}}{n!}+\frac{(\mathbf{D}-1)(n-1)}{4}+\sum_{\ell=1}^{n-2}(\sigma\sqrt{n})^{\ell}\Big(\binom{n-1}{\ell}\frac{2}{(\ell+1)!} \\
&+\frac{(\mathbf{D}-1)(\ell+1)}{4}\binom{n-1}{\ell+1}\frac{1}{(\ell+2)!}\Big)\Bigg)\cdot\frac{e^{-u/2}}{\sqrt{u}} \\
&\le \frac{\sqrt{\pi}}{2^{n-1}\Gamma(n/2)\Gamma((n+1)/2)}\big(\sqrt{2(\mathbf{D}-1)n}+\sqrt{\mathbf{D}u}\big)\sqrt{\mathbf{D}\mathcal{D}}\,n! \\
&\cdot\left(\sum_{\ell=0}^{n-1}\binom{n-1}{\ell}(\sigma\sqrt{n})^{\ell}+\frac{(\mathbf{D}-1)(n-1)}{4}\sum_{\ell=0}^{n-2}\binom{n-2}{\ell}(\sigma\sqrt{n})^{\ell}\right)\cdot\frac{e^{-u/2}}{\sqrt{u}}.
\end{aligned}
\tag{33}
$$

Now, we assume $n \ge 3$ and we bound this expectation for $0 < u < 1/(4\mathbf{D}^2n^5)$ in which case, by the bound for σ^2 given in (22), $\sigma^2 \le 4\mathbf{D}(\mathbf{D}-1)u \le 1/n^5$.
We will use throughout the bounds $1+x \le e^x$ for any x and $e^x - 1 \le 2x$ for $0 \le x \le 1$.

The factorial term $n! = \Gamma(n+1)$ and the other Gamma functions in the first line of the right-hand side of Inequality (33) can be bounded through Stirling's formula [1, Formula 6.1.38]: for any $x > 0$,

$$\Gamma(x+1) = \sqrt{2\pi x}\Big(\frac{x}{e}\Big)^x e^{\theta/(12x)} \quad \text{for some } 0 < \theta = \theta(x) < 1.$$

so that,

$$\sqrt{2\pi x}\Big(\frac{x}{e}\Big)^x < \Gamma(x+1) < \sqrt{2\pi x}\Big(\frac{x}{e}\Big)^x e^{1/(12x)}.$$

Also,

$$\sqrt{\mathbf{D}u} \le \frac{1}{2\sqrt{\mathbf{D}}n^{5/2}} \le \sqrt{2(\mathbf{D}-1)n}\Big(\frac{1}{2\sqrt{2(\mathbf{D}-1)n}\,\sqrt{\mathbf{D}}n^{5/2}}\Big) \le \frac{\sqrt{2(\mathbf{D}-1)n}}{4n^3},$$

which implies

$$\sqrt{2(\mathbf{D}-1)n}+\sqrt{\mathbf{D}u} \le \sqrt{2(\mathbf{D}-1)n}\,\Big(1+\frac{1}{4n^3}\Big).$$

Therefore, the first line of the right-hand side of Inequality (33) satisfies

$$\begin{aligned}
&\frac{\sqrt{\pi}}{2^{n-1}\Gamma(n/2)\Gamma((n+1)/2)}\left(\sqrt{2(\mathbf{D}-1)n}+\sqrt{\mathbf{D}u}\right)\sqrt{\mathbf{D}\mathcal{D}}\,n! \\
\leq\;& \frac{\sqrt{\pi}}{2^{n-1}\sqrt{(n-2)\pi}\sqrt{(n-1)\pi}}\left(\frac{2e}{n-2}\right)^{\frac{n-2}{2}}\left(\frac{2e}{n-1}\right)^{\frac{n-1}{2}}\sqrt{2(\mathbf{D}-1)n}\left(1+\frac{1}{4n^3}\right) \\
&\sqrt{\mathbf{D}\mathcal{D}}\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/(12n)} \\
=\;& \frac{e^{1/(12n)}}{e^{3/2}}\sqrt{\frac{n^2}{(n-2)(n-1)}}\left(\frac{n}{n-2}\right)^{\frac{n-2}{2}}\left(\frac{n}{n-1}\right)^{\frac{n-1}{2}} n^{3/2}\sqrt{2(\mathbf{D}-1)}\left(1+\frac{1}{4n^3}\right)\sqrt{\mathbf{D}\mathcal{D}} \\
\leq\;& 3\,\mathbf{D}\sqrt{\mathcal{D}}\,n^{3/2}\left(1+\frac{1}{4n^3}\right)\left(1+\frac{1}{6n}\right)\;\leq 4\,\mathbf{D}\sqrt{\mathcal{D}}\,n^{3/2}.
\end{aligned}$$

We now turn our attention to the term under brackets in the right-hand side of Inequality (33).
We have $\sigma\sqrt{n}\leq 1/n^2$. Therefore

$$\begin{aligned}
&\sum_{\ell=0}^{n-1}\binom{n-1}{\ell}(\sigma\sqrt{n})^\ell+\frac{(\mathbf{D}-1)(n-1)}{4}\sum_{\ell=0}^{n-2}\binom{n-2}{\ell}(\sigma\sqrt{n})^\ell \\
\leq\;& \left(1+\frac{1}{n^2}\right)^{n-1}+\frac{(\mathbf{D}-1)(n-1)}{4}\left(1+\frac{1}{n^2}\right)^{n-2} \\
\leq\;& e^{\frac{n-2}{n^2}}\left(1+\frac{1}{n^2}+\frac{(\mathbf{D}-1)(n-1)}{4}\right)\;\leq\;\left(1+\frac{2(n-2)}{n^2}\right)\left(1+\frac{1}{n^2}+\frac{(\mathbf{D}-1)(n-1)}{4}\right)\;\leq n\,\mathbf{D}.
\end{aligned}$$

Adding up, since $e^{-u/2}\leq 1$, we obtain

$$p_{\underline{L}}(u)\;\leq\;\mathbb{E}\big(H(u,\zeta)\big)\;\leq\;4\,\mathbf{D}^2\,\mathcal{D}^{1/2}\,n^{5/2}\,\frac{1}{\sqrt{u}}.$$

Step 7. We finally complete the proof of Theorem 1.2.

For $0<\alpha<1/(4\mathbf{D}^2n^5)$, the previous estimate for $p_{\underline{L}}(u)$ implies

$$\mathbb{P}(\underline{L}<\alpha)=\int_0^\alpha p_{\underline{L}}(u)\,du\;\leq 8\,\mathbf{D}^2\,\mathcal{D}^{1/2}\,n^{5/2}\sqrt{\alpha}.$$

Let us go back to the starting inequality (2):

$$\mathbb{P}\left(\widetilde{\kappa}(f)>a\right)\leq\mathbb{P}\left(\underline{L}<\frac{1}{a^2}(1+\ln a)N\right)+\exp\left(-\frac{N}{2}(\ln a-\ln(\ln a+1)\right).$$

where we recall that

$$N=\sum_{i=1}^n\binom{n+d_i}{n}\;\leq\;n^{\mathbf{D}+2}.$$

By hypothesis in the theorem, $a>a_n:=4\,\mathbf{D}^2n^3N^{1/2}$.
We set $\alpha:=(1+\ln a)N/a^2$ and verify $\alpha<1/(4\mathbf{D}^2n^5)$. It is enough to verify it with a_n:

$$\frac{(1+\ln a_n)N}{a_n^2}<\frac{1}{4\mathbf{D}^2n^5}\iff 1+\ln a_n\leq 4\mathbf{D}^2 n$$

which is satisfied since for $\mathbf{D} \geq 2$ and $n \geq 3$,

$$1+\ln a_n < 1+\ln(4\mathbf{D}^2)+3\ln n+\frac{\mathbf{D}+2}{2}\ln n \leq 4\mathbf{D}^2+\big(\frac{\mathbf{D}}{2}+4\big)\ln n < 4\mathbf{D}^2+4\mathbf{D}^2(n-1) = 4\mathbf{D}^2 n.$$

Therefore, by Inequality (2),

$$\begin{aligned}\mathbb{P}(\widetilde{\kappa}(f)) > a) \leq\ & \mathbb{P}\Big(\underline{L} < \frac{(1+\ln a)N}{a^2}\Big) + \exp\Big(-\frac{N}{2}\big(\ln a - \ln(\ln a+1)\big)\Big) \\ \leq\ & 8\,\mathbf{D}^2\,\mathcal{D}^{1/2}\,n^{5/2}\,\sqrt{\alpha} + \frac{1}{a} \ = \ K_n\frac{(1+\ln a)^{1/2}}{a}\end{aligned}$$

where $K_n = 8\mathbf{D}^2\mathcal{D}^{1/2}\,N^{1/2}n^{5/2}+1$. Here we used $\exp\big((-N/2)\big(\ln a - \ln(\ln a+1)\big)\big) < 1/a$ for $a > 2, N > 10$. This proves part (i) of Theorem 1.2.

(ii) We verify that $K_n > a_n$. It is enough to check

$$8\,\mathbf{D}^2\mathcal{D}^{1/2}N^{1/2}n^{5/2} \geq 4\,\mathbf{D}^2 n^3 N^{1/2} \iff 2\,\mathcal{D}^{1/2} \geq n^{1/2} \iff 4\,\mathcal{D} \geq n$$

which holds because $4\,\mathcal{D} \geq 4\cdot 2^n \geq n$.
Therefore we can write

$$\begin{aligned}\mathbb{E}(\ln\widetilde{\kappa}(f)) =\ & \int_0^{+\infty}\mathbb{P}(\ln\widetilde{\kappa}(f) > x)\,dx \ \leq\ \ln K_n + \int_{\ln K_n}^{+\infty}\mathbb{P}\big(\widetilde{\kappa}(f) > e^x\big)\,dx \\ \leq\ & \ln K_n + \int_{\ln K_n}^{+\infty} K_n(1+x)^{1/2}e^{-x}\,dx \\ \leq\ & \ln K_n + K_n\int_{\ln K_n}^{+\infty} x^{1/2}e^{-x}\,dx + \frac{K_n}{2}\int_{\ln K_n}^{+\infty} x^{-1/2}e^{-x}\,dx \\ =\ & \ln K_n + K_n(e^{-\ln K_n}\,(\ln K_n)^{1/2}) + K_n\int_{\ln K_n}^{+\infty} x^{-1/2}e^{-x}\,dx \\ \leq\ & \ln K_n + (\ln K_n)^{1/2} + K_n(\ln K_n)^{-1/2}\int_{\ln K_n}^{+\infty} e^{-x}\,dx \\ =\ & \ln K_n + (\ln K_n)^{1/2} + (\ln K_n)^{-1/2}.\end{aligned}$$

Here we used the inequality $(1+x)^{1/2} < x^{1/2} + \frac{1}{2}x^{-1/2}$ for $x > 0$ and integration by parts.

3 Auxiliary lemmas

This section contains the proofs of all the auxiliary results indicated by the symbol $\diamondsuit$, which were stated without proof during the text.

Proof of Lemma 2.1. According to the definition of the Weyl norm,

$$\|f\|_W^2 = \sum_{i=1}^{n}\sum_{|j|=d_i}\xi_{i,j}^2 \tag{34}$$

where, due to the distribution, the random variables

$$\xi_{i,j} = \frac{a_j^{(i)}}{\binom{d_i}{j}^{1/2}}$$

are independent identically distributed (i.i.d.) standard normal.
It is easy to see that the number of terms in the sum (34) is equal to N, so that

$$\mathbb{P}\left(\|f\|_W^2 \geq (1+\eta)N\right) = \mathbb{P}\left((\xi_1^2 - 1) + \cdots + (\xi_N^2 - 1) \geq \eta N\right) = \mathbb{P}\left(\frac{X_1 + \ldots + X_N}{N} \geq \eta\right)$$

where $X_1, \ldots, X_N$ are i.i.d. random variables having the distribution of $\xi^2 - 1$, ξ a normal standard random variable.
The logarithmic moment generating function of $\xi^2 - 1$ is

$$\Lambda(\lambda) = \ln \mathbb{E}\{e^{\lambda(\xi^2-1)}\} = \begin{cases} -\lambda - \frac{1}{2}\ln(1-2\lambda) & \text{if } \lambda < \frac{1}{2} \\ +\infty & \text{if } \lambda \geq \frac{1}{2} \end{cases}$$

and its Fenchel-Legendre transform

$$\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}}(\lambda x - \Lambda(\lambda)) = \begin{cases} \frac{1}{2}(x - \ln(x+1)) & \text{if } x > -1 \\ +\infty & \text{if } x \leq -1. \end{cases}$$

A basic result on large deviations [11, Ch. 2] states that, for any integer m and any $x > 0$,

$$\mathbb{P}\left(\frac{X_1 + \cdots + X_m}{m} \geq x\right) \leq \exp(-m\Lambda^*(x)).$$

This implies the statement. $\square$

Proof of Lemma 2.2. For the first item, from the fact that $\mathbb{E}(a_j a_{j'}) = \mathbb{E}(a_j)\mathbb{E}(a_{j'}) = 0$ for $j \neq j'$ (by the independence of the a_j), we have

$$\mathbb{E}(f(x)f(y)) = \mathbb{E}\left(\sum_{j,j'} a_j a_{j'} x^j y^{j'}\right) = \sum_j \mathbb{E}((a_j)^2) x^j y^j = \sum_j \binom{d}{j} x^j y^j = \langle x, y\rangle^d.$$

For the following items, we observe that we can differentiate under the expectation sign the function $(x,y) \mapsto \mathbb{E}(f(x)f(y)) = \langle x, y\rangle^d$, e.g.

$$\begin{aligned}
\mathbb{E}\left(f(x)\partial_k f(y)\right) &= \frac{\partial(\langle x,y\rangle^d)}{\partial y_k}(x,y) = d x_k \langle x,y\rangle^{d-1} \\
\mathbb{E}\left(\partial_k f(x)\partial_{k'} f(y)\right) &= \partial^2_{kk'}(\langle x,y\rangle^d) = \delta_{kk'} d\langle x,y\rangle^{d-1} + d(d-1)x_{k'}y_k\langle x,y\rangle^{d-2}.
\end{aligned}$$

This gives the covariances when specializing $x = y = e_0$. $\square$

Our next lemma deals with the analytic description of the geometry of the manifold V which is used in the proof of Lemma 2.3. We define the function $\psi : B_{2n-1,\delta} \to \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ by means of:

$$\psi(\sigma_2, \ldots, \sigma_n, \tau_2, \ldots, \tau_n, \theta) = \left(\frac{C}{\|C\|_{n+1}}, \frac{D}{\|D\|_{n+1}}\right),$$

where $B_{2n-1,\delta}$ is the open ball in $\mathbb{R}^{2n-1}$, centered at the origin and radius δ sufficiently small, $\|.\|_{n+1}$ is the Euclidean norm in $\mathbb{R}^{n+1}$ and the definition of C and D is given in several steps by the following:

- We set $\sigma_1 := (1-\sigma_2^2 - \ldots - \sigma_n^2)^{1/2},\ \tau_1 := (1-\tau_2^2 - \ldots - \tau_n^2)^{1/2}$,

  $a(\sigma,\tau) := -\left(\sum_{j=2}^n \sigma_j\tau_j\right)/(\sigma_1+\tau_1), \quad n(\sigma,\tau) := \sqrt{1+a^2(\sigma,\tau)}.$
- $A := \frac{1}{n(\sigma,\tau)}\left(\sigma_1 e_0 + \sum_{j=2}^n \sigma_j e_j + a(\sigma,\tau)e_1\right)$, and

  $B := \frac{1}{n(\sigma,\tau)}\left(\tau_1 e_1 + \sum_{j=2}^n \tau_j e_j + a(\sigma,\tau)e_0\right)$.
- $C := \cos(\theta/\sqrt{2})A + \sin(\theta/\sqrt{2})\sigma_1 e_1$, and $D := \cos(\theta/\sqrt{2})B - \sin(\theta/\sqrt{2})\tau_1 e_0$.

Lemma 3.1. ***[Geometry of*** V***]***

1. *ψ is a parametrization of a neighborhood of the point (e_0,e_1) in the manifold V with $\psi(0)=(e_0,e_1)$.*
2. *For $2\le j\le n$,*

$$\frac{\partial\psi}{\partial\sigma_j}(0) = (e_j,0), \quad \frac{\partial\psi}{\partial\tau_j}(0) = (0,e_j) \quad \text{and} \quad \frac{\partial\psi}{\partial\theta}(0) = \frac{1}{\sqrt{2}}(e_1,-e_0).$$

Therefore the orthonormal basis $\mathcal{B}_T$ (defined in (12)) of the tangent space of V at the point (e_0,e_1) satisfies

$$\mathcal{B}_T = \left(\frac{\partial\psi}{\partial\sigma_2}(0),\ldots,\frac{\partial\psi}{\partial\sigma_n}(0),\frac{\partial\psi}{\partial\tau_2}(0),\ldots,\frac{\partial\psi}{\partial\tau_n}(0),\frac{\partial\psi}{\partial\theta}(0)\right).$$

3. *The curvatures are given by:*

$$\begin{aligned}
&\frac{\partial^2\psi}{\partial\sigma_j^2}(0) = (-e_0,0);\ \frac{\partial^2\psi}{\partial\tau_j^2} = (0,-e_1);\ \frac{\partial^2\psi}{\partial\sigma_j\partial\tau_j} = -\frac{1}{2}(e_1,e_0) \quad \text{for } 2\le j\le n,\\
&\frac{\partial^2\psi}{\partial\sigma_j\partial\sigma_k}(0) = \frac{\partial^2\psi}{\partial\tau_j\partial\tau_k}(0) = \frac{\partial^2\psi}{\partial\sigma_j\partial\tau_k}(0) = (0,0) \quad \text{for } 2\le j\ne k\le n,\\
&\frac{\partial^2\psi}{\partial\theta^2}(0) = -\frac{1}{2}(e_0,e_1);\ \frac{\partial^2\psi}{\partial\sigma_j\partial\theta}(0) = \frac{\partial^2\psi}{\partial\tau_j\partial\theta} = (0,0) \quad \text{for } 2\le j\le n.
\end{aligned}$$

PROOF. If δ is small enough, ψ is well defined and is $\mathscr{C}^\infty$. It is easy to check that $\langle C,D\rangle_{\mathbb{R}^{n+1}} = 0$, so that $\psi(\sigma_2,\ldots,\sigma_n,\tau_2,\ldots,\tau_n,\theta)\in V$.
A routine calculation of first derivatives allows to check *2* and also implies that if δ is small enough, ψ is a diffeomorphism from $B(0,\delta)$ onto its image. The computation of second order derivatives is also immediate. □

Corollary 3.2. *Let us set $L' := L'(e_0,e_1)$ and $L'' := L''(e_0,e_1)$ for the free first order and second order derivatives of L at (e_0,e_1). We use the parametrization introduced in the previous Lemma. Consider the function*

$$\widetilde{L}(\sigma_2,\ldots,\sigma_n,\tau_2,\ldots,\tau_n,\theta) = L\big(\psi(\sigma_2,\ldots,\sigma_n,\tau_2,\ldots,\tau_n,\theta)\big)$$

Let M be the symmetric matrix of the linear operator $\widetilde{L}''(0)$ in the canonical basis of $\mathbb{R}^{2n-1}$*:*

$$M = \begin{pmatrix} M_{\sigma\sigma} & M_{\sigma\tau} & M_{\sigma\theta} \\ M_{\tau\sigma} & M_{\tau\tau} & M_{\tau\theta} \\ M_{\theta\sigma} & M_{\theta\tau} & M_{\theta\theta} \end{pmatrix} \in \mathbb{R}^{(2n-1)\times(2n-1)}$$

where for $2 \leq j, k \leq n$,

$$\begin{aligned}
(M_{\sigma\sigma})_{jk} = (M_{\sigma\sigma})_{kj} &= \frac{\partial^2 (L\circ\psi)}{\partial\sigma_j \partial\sigma_k} = \left\langle L'' \frac{\partial\psi}{\partial\sigma_j}(0), \frac{\partial\psi}{\partial\sigma_k}(0)\right\rangle + \left\langle L', \frac{\partial^2\psi}{\partial\sigma_j\partial\sigma_k}(0)\right\rangle \\
&= \begin{cases} \langle L''(e_j,0),(e_j,0)\rangle - \langle L',(e_0,0)\rangle & \textit{for} \quad j=k \\ \langle L''(e_j,0),(e_k,0)\rangle & \textit{for} \quad j\neq k \end{cases} \\
&= \begin{cases} \frac{\partial^2 L}{\partial x_j^2} - \frac{\partial L}{\partial x_0} & \textit{for} \quad j=k \\ \frac{\partial^2 L}{\partial x_j \partial x_k} & \textit{for} \quad j\neq k \end{cases}
\end{aligned}$$

$$\begin{aligned}
(M_{\sigma\tau})_{jk} = (M_{\tau\sigma})_{kj} &= \frac{\partial^2 (L\circ\psi)}{\partial\sigma_j \partial\tau_k} = \left\langle L'' \frac{\partial\psi}{\partial\sigma_j}(0), \frac{\partial\psi}{\partial\tau_k}(0)\right\rangle + \left\langle L', \frac{\partial^2\psi}{\partial\sigma_j\partial\tau_k}(0)\right\rangle \\
&= \begin{cases} \langle L''(e_j,0),(0,e_j)\rangle - \frac{1}{2}\langle L',(e_1,e_0)\rangle & \textit{for} \quad j=k \\ \langle L''(e_j,0),(0,e_k)\rangle & \textit{for} \quad j\neq k \end{cases} \\
&= \begin{cases} \frac{\partial^2 L}{\partial x_j \partial y_j} - \frac{1}{2}(\frac{\partial L}{\partial x_1} + \frac{\partial L}{\partial y_0}) & \textit{for} \quad j=k \\ \frac{\partial^2 L}{\partial x_j \partial y_k} & \textit{for} \quad j\neq k \end{cases}
\end{aligned}$$

$$\begin{aligned}
(M_{\tau\tau})_{jk} = (M_{\tau\tau})_{kj} &= \frac{\partial^2 (L\circ\psi)}{\partial\tau_j \partial\tau_k} = \left\langle L'' \frac{\partial\psi}{\partial\tau_j}(0), \frac{\partial\psi}{\partial\tau_k}(0)\right\rangle + \left\langle L', \frac{\partial^2\psi}{\partial\tau_j\partial\tau_k}(0)\right\rangle \\
&= \begin{cases} \langle L''(0,e_j),(0,e_j)\rangle - \langle L',(0,e_1)\rangle & \textit{for} \quad j=k \\ \langle L''(0,e_j),(0,e_k)\rangle & \textit{for} \quad j\neq k \end{cases} \\
&= \begin{cases} \frac{\partial^2 L}{\partial y_j^2} - \frac{\partial L}{\partial y_1} & \textit{for} \quad j=k \\ \frac{\partial^2 L}{\partial y_j \partial y_k} & \textit{for} \quad j\neq k \end{cases},
\end{aligned}$$

for $2 \leq j \leq n$,

$$\begin{aligned}
(M_{\sigma\theta})_{j1} = (M_{\theta\sigma})_{1j} &= \frac{\partial^2 (L\circ\psi)}{\partial\sigma_j \partial\theta} = \left\langle L'' \frac{\partial\psi}{\partial\sigma_j}(0), \frac{\partial\psi}{\partial\theta}(0)\right\rangle + \left\langle L', \frac{\partial^2\psi}{\partial\sigma_j\partial\theta}(0)\right\rangle \\
&= \frac{1}{\sqrt{2}}\langle L''(e_j,0),(e_1,-e_0)\rangle \;=\; \frac{1}{\sqrt{2}}\left(\frac{\partial^2 L}{\partial x_j\partial x_1} - \frac{\partial^2 L}{\partial x_j \partial y_0}\right),
\end{aligned}$$

$$\begin{aligned}
(M_{\tau\theta})_{j1} = (M_{\theta\tau})_{1j} &= \frac{\partial^2 (L\circ\psi)}{\partial\tau_j \partial\theta} = \left\langle L'' \frac{\partial\psi}{\partial\tau_j}(0), \frac{\partial\psi}{\partial\theta}(0)\right\rangle + \left\langle L', \frac{\partial^2\psi}{\partial\tau_j\partial\theta}(0)\right\rangle \\
&= \frac{1}{\sqrt{2}}\langle L''(0,e_j),(e_1,-e_0)\rangle \;=\; \frac{1}{\sqrt{2}}\left(\frac{\partial^2 L}{\partial y_j\partial x_1} - \frac{\partial^2 L}{\partial y_j \partial y_0}\right),
\end{aligned}$$

and finally

$$
\begin{aligned}
M_{\theta\theta} &= \frac{\partial^2 (L\circ\psi)}{\partial\theta^2} \;=\; \left\langle L''\frac{\partial\psi}{\partial\theta}(0), \frac{\partial\psi}{\partial\theta}(0)\right\rangle + \left\langle L', \frac{\partial^2\psi}{\partial\theta^2}(0)\right\rangle \\
&= \frac{1}{2}\left(\langle L''(e_1,-e_0),(e_1,-e_0)\rangle - \langle L',(e_0,e_1)\rangle\right) \\
&= \frac{1}{2}\left(\frac{\partial^2 L}{\partial x_1^2} - 2\frac{\partial^2 L}{\partial x_1\partial y_0} + \frac{\partial^2 L}{\partial y_0^2} - \frac{\partial L}{\partial x_0} - \frac{\partial L}{\partial y_1}\right).
\end{aligned}
$$

□

PROOF OF LEMMA 2.4. We have:

$$
\mathbb{E}\big(g(X)\,/\,XY+Z=u, Y=0\big) = \int_{\mathbb{R}^{p\times q}} g(x)\frac{p_{X,Y,XY+Z}(x,0,u)}{p_{Y,XY+Z}(0,u)}\;dx \tag{35}
$$

since

$$
\frac{p_{X,Y,XY+Z}(x,0,u)}{p_{Y,XY+Z}(0,u)}
$$

is the conditional density of X at the point x, given that $Y=0, XY+Z=u$.
Now, the density $p_{X,Y,XY+Z}(x,y,u)$ is easily computed from the change of variables formula (using the independence of $X,\ Y,\ Z$), obtaining:

$$
p_{X,Y,XY+Z}(x,y,u) = p_X(x)p_Y(y)p_Z(u-xy).
$$

This also implies

$$
p_{Y,XY+Z}(0,u) = \int_{\mathbb{R}^{p\times q}} p_{X,Y,XY+Z}(x,0,u)\;dx = p_Y(0)p_Z(u).
$$

Replacing $p_{Y,XY+Z}(0,u)$ by $p_Y(0)p_Z(u)$ in (35), we get:

$$
\mathbb{E}\big(g(X)\,/\,XY+Z=u, Y=0\big) = \int_{\mathbb{R}^{p\times q}} g(x)p_X(x)\;dx = \mathbb{E}\big(g(X)\big)
$$

□

PROOF OF LEMMA 2.5. Write the Taylor expansion of $\det(C_q(\lambda))$ at $\lambda=0$ and compute the successive derivatives at this point. □

PROOF OF LEMMA 2.6. We note that $Q = M\,M^t$ where

$$
M := \begin{array}{cc}
\begin{array}{cc} \scriptstyle n & \scriptstyle n \end{array} & \\
\left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right) & \begin{array}{l} \scriptstyle k \\ \scriptstyle n-1 \end{array}
\end{array}.
$$

Applying the Cauchy-Binet formula (29), we get

$$
\det(Q) = \sum_{\#(S')=k+n-1} \big(\det(M^{S'})\big)^2, \tag{36}
$$

where the sum is over all choices of $k+n-1$ columns of M.

We fix such an S'. It is easy to see, performing a Laplace expansion with respect to the first k rows of the obtained matrix, that if we take strictly more than k columns in the n-columns right block corresponding to B, then $\det(M^{S'}) = 0$. This is because in this expansion there will always remain a zero column. Therefore, we can only choose up to k columns in the right block, i.e. there are two cases: we choose all the n columns in the left block and $k-1$ columns in the right block, or we choose $n-1$ columns in the left block and k columns in the right block.

Case 1: $M^{S'}$ is of the form:

$$M^{S'} = \begin{array}{cc} \begin{array}{cc} \scriptstyle n & \scriptstyle k-1 \end{array} & \\ \left(\begin{array}{c|c} A & B^S \\ \hline C & 0 \end{array}\right) & \begin{array}{c} \scriptstyle k \\ \scriptstyle n-1 \end{array} \end{array} \in \mathbb{R}^{(k+n-1)\times(k+n-1)}.$$

Here S is the set of $(k-1)$ columns of B that we kept. Again using Laplace expansion with respect to the last $k-1$ columns of $M^{S'}$, we see that each non-zero determinant corresponds to suppressing a row –say row i– of B^S, times the determinant of its complementary matrix which is equal to the i-th row of A added to C. Finally, expanding this last matrix by the i-th row of A, we obtain:

$$\det(M^{S'}) = (-1)^{n(k-1)} \sum_{i=1}^{k} (-1)^{k-i} \det(B_{\overline{i}}^S \sum_{j=1}^{n} (-1)^{j-1} a_{ij} \det(C^{\overline{j}}),$$

where $\overline{i}$ and $\overline{j}$ denote the complementary rows or columns, accordingly.

Case 2: $M^{S'}$ is of the following form for some j which corresponds to the suppressed column of A and S is a choice of k columns of B:

$$M^{S'} = \begin{array}{cc} \begin{array}{cc} \scriptstyle n-1 & \scriptstyle k \end{array} & \\ \left(\begin{array}{c|c} A^{\overline{j}} & B^S \\ \hline C^{\overline{j}} & 0 \end{array}\right) & \begin{array}{c} \scriptstyle k \\ \scriptstyle n-1 \end{array} \end{array} \in \mathbb{R}^{(k+n-1)\times(k+n-1)}.$$

Then, permuting the two blocks of rows and since the obtained matrix is block-diagonal, we get $\det(M^S) = (-1)^{k(n-1)} \det(C^{\overline{j}}) \det(B^S)$.

Therefore, the sum in (36) for all S' in Case 2 gives:

$$\sum_{j=1}^{n} \left(\det(C^{\overline{j}})\right)^2 \sum_{\#(S)=k} \left(\det(B^S)\right)^2 = \det(CC^t)\det(BB^t),$$

again by the Cauchy-Binet formula (29). The statement follows from adding up over all S' in Cases 1 and 2. □

Proof of Lemma 2.7. The proof is based on the following bound for the tails of the probability distribution of $\overline{\lambda}$. For $t > 0$ one has (see for example [10] and references therein):

$$\mathbb{P}(\overline{\lambda} \geq 2 + \sqrt{2}\ t) < \exp\big(-\frac{nt^2}{2}\big).$$

Therefore, since $\ell \leq n$,

$$\begin{aligned}
\mathbb{E}\big(\overline{\lambda}^{\ell}\big) &= \int_0^{+\infty} \mathbb{P}(\overline{\lambda}^{\ell} > x)\, dx \ = \int_0^{+\infty} \mathbb{P}(\overline{\lambda} > y)\ \ell y^{\ell-1}\, dy \\
&\leq \int_0^4 \ell y^{\ell-1}\, dy \ + \int_4^{+\infty} \ell\, y^{\ell-1} \exp\Big(-\frac{n}{2}\cdot\frac{(y-2)^2}{2}\Big)\, dy \\
&\leq 4^{\ell} \ + \int_{\sqrt{2n}}^{+\infty} \ell\sqrt{\frac{2}{n}}\Big(\sqrt{\frac{2}{n}}u + 2\Big)^{\ell-1} \exp\Big(-\frac{u^2}{2}\Big)\, du \\
&= 4^{\ell} \ + \ell\sqrt{\frac{2}{n}}\, 2^{\ell-1} \int_{\sqrt{2n}}^{+\infty} \Big(1 + \frac{u}{\sqrt{2n}}\Big)^{\ell-1} \exp\Big(-\frac{u^2}{2}\Big)\, du \\
&\leq 4^{\ell} + \ell\sqrt{\frac{2}{n}}\, 2^{\ell-1} \int_{\sqrt{2n}}^{+\infty} \exp\big(u\sqrt{\frac{n}{2}} - \frac{u^2}{2}\big)\, du \quad \text{since} \quad 1 + x \leq \exp(x) \\
&= 4^{\ell} \ + \ell\sqrt{\frac{2}{n}}\, 2^{\ell-1} \exp(n/4) \int_{\sqrt{n/2}}^{+\infty} \exp\big(-\frac{y^2}{2}\big)\, dy \\
&\leq 4^{\ell} \ + \ell\sqrt{\frac{2}{n}}\, 2^{\ell-1} \exp\Big(\frac{n}{4}\Big)\sqrt{\frac{2}{n}} \exp\Big(-\frac{n}{4}\Big) \ \leq \ 2\cdot 4^{\ell},
\end{aligned}$$

where in the last line we used that

$$\int_a^{+\infty} \exp\Big(-\frac{y^2}{2}\Big)\, dy < \int_a^{+\infty} \frac{y}{a} \exp\Big(-\frac{y^2}{2}\Big)\, dy = \frac{1}{a}\exp\Big(-\frac{a^2}{2}\Big). \qquad \square$$

Acknowledgment. We are thankful to the anonymous referee for his many suggestions that helped us improving the presentation of this text.

References

[1] M. Abramowitz and I. Stegun, Handbook of mathematical functions with formulas, graphs, and mathematical tables, National Bureau of Standards Applied Mathematics Series, 55, 1964.

[2] J.-M. Azaïs, M. Wschebor, Level sets and extrema of random processes and fields, John Wiley and Sons, 2009.

[3] C. Beltrán, L.M. Pardo, On Smale's 17 problem: a probabilistic positive solution, Found. Comput. Math., 8 (2008) 1–43.

[4] C. Beltrán, M. Shub, Complexity of Bézout's theorem VII: distance estimates in the condition metric, Found. Comput. Math., 9 (2009) 179–195.

[5] C.E. Borges, L.M. Pardo, On the probability distribution of data at points in real complete intersection varieties, J. Compl., 24 (2008) 492–523.

[6] P. Bürgisser, F. Cucker, M. Lotz, The probability that a slightly perturbed numerical analysis problem is difficult, Mathematics of Computation, 77 (2008) 1559–1583.

[7] F. Cucker, T. Krick, G. Malajovich, M. Wschebor, A numerical algorithm for zero counting. I: Complexity and accuracy, J. Compl., 24 (2008) 582–605.

[8] F. Cucker, T. Krick, G. Malajovich, M. Wschebor, A numerical algorithm for zero counting. II: Distance to Ill-posedness and Smoothed Analysis, J. fixed point theory appl., 6 (2009) 285–294.

[9] F. Cucker, S. Smale, Complexity estimates depending on condition and round-off error, Journal of the ACM, 46 (1999) 113–184.

[10] K.R. Davidson, S.J. Szarek, Local Operator Theory, Random Matrices and Banach Spaces, in: W.B. Johnson, J. Lindenstrauss (Eds.), Handbook of the Geometry of Banach Spaces, North Holland, 2001, pp. 317-366.

[11] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Springer-Verlag, 2nd edition, 1998.

[12] J. Demmel, On condition numbers and the distance to the nearest ill-posed problem, Numer. Math., 51 (1987) 251–289.

[13] R.M. Freund, J.R. Vera, Some characterizations and properties of the "distance to ill-posedness" and the condition measure of a conic linear system, Math. Program., 86 (1999) 225–260.

[14] F.R. Gantmacher, The Theory of Matrices, AMS-Chelsea, 2000.

[15] E. Kostlan, Random polynomials and the statistical fundamental theorem of algebra, unpublished, 1987.

[16] S. Rump, Ill-conditioned matrices are componentwise near to singularity, SIAM Review, 41 (1999) 102–112.

[17] M. Shub, Complexity of Bézout's theorem VI: geodesics in the condition (number) metric, Found. Comput. Math., 9 (2009) 171–178.

[18] M. Shub, S. Smale, Complexity of Bézout's theorem I: geometric aspects, Journal of the Amer. Math. Soc., 6 (1993) 459–501.

[19] M. Shub, S. Smale, Complexity of Bézout's theorem II: volumes and probabilities, in: F. Eyssette and A. Galligo, editors, Computational Algebraic Geometry, volume 109 of Progress in Mathematics, pages 267–285. Birkhäuser, 1993.

[20] M. Shub, S. Smale, Complexity of Bézout's theorem III: condition number and packing, Journal of Complexity, 9 (1993) 4–14.

[21] M. Shub, S. Smale, Complexity of Bézout's theorem IV: probability of success; extensions, SIAM J. of Numer. Anal., 33 (1996) 128–148.

[22] M. Shub, S. Smale, Complexity of Bézout's theorem V: polynomial time, Theoret. Comp. Sci., 133 (1994) 141–164.

[23] J. Wilkinson, Note on matrices with a very ill-conditioned eigenproblem, Numer. Math., 19 (1972) 176–178.